



Department of Mathematics

# Four Generations of Asset Pricing Models and Volatility Dynamics

Diploma Thesis  
by Sascha Desmettre

supervised by  
Prof. Dr. Ralf Korn

submitted June 15, 2007



## Abstract

The scope of this diploma thesis is to examine the four generations of asset pricing models and the corresponding volatility dynamics which have been developed so far. We proceed as follows: In chapter 1 we give a short repetition of the Black-Scholes **first generation model** which assumes a constant volatility and we show that volatility should not be modeled as constant by examining statistical data and introducing the notion of implied volatility.

In chapter 2, we examine the simplest models that are able to produce smiles or skews - local volatility models. These are called **second generation models**. Local volatility models model the volatility as a function of the stock price and time. We start with the work of Dupire, show how local volatility models can be calibrated and end with a detailed discussion of the constant elasticity of volatility model.

Chapter 3 focuses on the Heston model which represents the class of the stochastic volatility models, which assume that the volatility itself is driven by a stochastic process. These are called **third generation models**. We introduce the model structure, derive a partial differential pricing equation, give a closed-form solution for European calls by solving this equation and explain how the model is calibrated. The last part of chapter 3 then deals with the limits and the mis-specifications of the Heston model, in particular for recent exotic options like reverse cliquets, Accumulators or Napoleons.

In chapter 4 we then introduce the Bergomi forward variance model which is called **fourth generation model** as a consequence of the limits of the Heston model explained in chapter 3. The Bergomi model is a stochastic local volatility model - the spot price is modeled as a constant elasticity of volatility diffusion and its volatility parameters are functions of the so called forward variances which are specified as stochastic processes. We start with the model specification, derive a partial differential pricing equation, show how the model has to be calibrated and end with pricing examples and a concluding discussion.

## Acknowledgements

At first, I want to thank Prof. Dr. Ralf Korn for giving me the opportunity to deal with this interesting topic and for his competent comments on my work. Discussions with him were always very helpful and inspiring.

Further I want to thank Dr. Ulrich Nögel who always was willing to answer my questions and gave me some useful hints and Dr. Lorenzo Bergomi who kindly answered my questions concerning his work on the Bergomi forward variance model.

Moreover, I want to take the opportunity to express my deep gratitude to my parents and my sister for supporting me over the years.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>i</b>
<b>1 Black-Scholes Framework</b>	<b>1</b>
1.1 Model Structure . . . . .	1
1.2 Some stylized Facts and implied Volatility . . . . .	6
Digression 1: Volatility Smiles, Skews and Frowns . . . . .	14
<b>2 Local Volatility Models</b>	<b>16</b>
2.1 The Formula of Dupire . . . . .	16
2.2 Calibration: A practical Approach . . . . .	20
Digression 2: Calibrating a Model . . . . .	20
2.3 The CEV Model as a local Volatility Model . . . . .	24
2.3.1 Weak Solutions by Squared Bessel Processes . . . . .	25
2.3.2 Existence of a unique equivalent Martingale Measure . . . . .	28
2.3.3 Closed-form Solution for European Calls . . . . .	32
<b>3 Heston and its Shortfalls</b>	<b>42</b>
3.1 Motivation . . . . .	42
3.2 Model Structure . . . . .	43
3.3 Derivation of the Pricing Equation . . . . .	45
3.4 Closed-form Solution for European Calls . . . . .	51
Digression 3: Change of Numéraire . . . . .	52
Digression 4: Characteristic Functions and Distribution Functions . . . . .	59
3.5 Calibration and numerical Results . . . . .	68
3.6 The Limits of Heston . . . . .	70
3.6.1 Example: Napoleon and its Dependence on Future Smiles . . . . .	70
3.6.2 Static Properties of the Heston Model . . . . .	72
3.6.3 Dynamics of the Heston Parameters . . . . .	75
3.6.4 Dynamics of implied Volatilities . . . . .	77
3.6.5 Forward-starting Options . . . . .	79
3.6.6 Local Dynamics . . . . .	81
<b>4 The Bergomi Forward Variance Model</b>	<b>84</b>
4.1 Dynamics for the Forward Variances . . . . .	85
4.1.1 A One-factor Model . . . . .	86
4.1.2 A Two-factor Model . . . . .	89
4.1.3 A discrete LIBOR-type Structure . . . . .	91
4.1.4 An N-factor Model . . . . .	93
4.2 Joint Dynamics for the Spot Process . . . . .	93
4.2.1 A continuous Setting . . . . .	93
4.2.2 A discrete Setting . . . . .	94
4.3 The Pricing Equation . . . . .	96
4.3.1 The 2-factor Model . . . . .	96
4.3.2 The N-factor Model . . . . .	98
4.4 Three-step Calibration . . . . .	99
4.4.1 Set Dynamics for the implied Variance Swap Volatilities . . . . .	100

4.4.2	Calibration of the short Forward Skew . . . . .	102
4.4.3	Setting Correlations between the Spot and short/long Factors - The Term Skew . . . . .	103
4.5	Pricing Examples . . . . .	106
4.6	Summary/Open Questions . . . . .	109
<b>References</b>		<b>110</b>

## List of Figures

1	DAX daily Log Returns from 1990-1992 . . . . .	7
2	Frequency Distribution of DAX daily Log Returns . . . . .	8
3	Normplot of DAX daily Log Returns . . . . .	8
4	Volatility Surface of S&P 500 and N225 Index . . . . .	13
5	Volatility Surface assumed by Black and Scholes . . . . .	13
6	Scheme of a Volatility Smile . . . . .	14
7	Volatility Skew of S&P 500 and N225 Index . . . . .	15
8	Scheme of a Volatility Frown . . . . .	15
9	Rolling one-year historic Volatility of S&P 500 Index and N225 Index . . . . .	43
10	Integration Path $C_{M,\epsilon}$ . . . . .	64
11	S&P 500 Volatility Surface and relative Errors after Calibration . . . . .	69
12	Price and Vega of Napoleon . . . . .	72
13	Variance Swap Volatility and 1-Year ATM Volatility . . . . .	75
14	Time Evolution of Heston Parameters for SPX Index . . . . .	75
15	Quantitative Stock-Volatility Correlation and Time Evolution of $\rho$ . . . . .	76
16	Time Evolution of Vol-of-Vol compared to Start Vol . . . . .	77
17	Running monthly Averages of $\nu_{\frac{\delta S}{S}}^{\text{real}}$ and $\nu_{\frac{\delta S}{S}}^{\text{impl}}$ , $\nu_{\delta\nu}^{\text{real}}$ and $\nu_{\delta\nu}^{\text{impl}}$ , $C_{\frac{\delta S}{S}\delta\nu}^{\text{real}}$ and $C_{\frac{\delta S}{S}\delta\nu}^{\text{impl}}$ . . . . .	78
18	Forward Smile for 1-Year and 3-Month Maturity . . . . .	81
19	Tenor Structure of Forward Variances . . . . .	92
20	Illustration for the Definition of $\sigma_{Vol}(\tau)$ . . . . .	100
21	Term Structure of $\sigma_{Vol}(\tau)$ for a One-month Interval . . . . .	101
22	Term Structure of $\sigma_{Vol}(\tau)$ for a One-year Interval . . . . .	101
23	$\sigma_0 \left( \hat{\sigma}_{VS}^{(i)} \right)$ and $\beta \left( \hat{\sigma}_{VS}^{(i)} \right)$ for constant Skew and proportional Skew . . . . .	102
24	Approximate Skew Graphs . . . . .	105

## List of Tables

1	Sample Kurtosis of DAX Shares and DAX . . . . .	10
2	Sample Skewness of DAX Shares and DAX . . . . .	11
3	Prices of the Bergomi Model with constant Skew . . . . .	107
4	Prices of the Bergomi Model with constant Skew and Correlations halved . . . . .	108
5	Prices of the Bergomi Model with proportional Skew . . . . .	109

# 1 Black-Scholes Framework

The model developed by Black and Scholes is the so called **first generation model** and had a deep impact on the financial market when it came out 1973. In this chapter we will introduce the model structure and derive the well-known PDE of Black and Scholes and provide as well a solution of it. Then, we will give evidence for the fact that the assumption of a constant volatility in the Black-Scholes model is doubtful and so motivate why a development of more advanced equity price models is necessary.

## 1.1 Model Structure

Principally, we consider a continuous-time security market consisting of an interest rate-based cash account (money market account(MMA)) and  $n$  risky assets ( $i = 1, \dots, n$ ):

$$dB(t) = B(t)r(t)dt \quad ; \quad B(0) = 1 \quad (1.1.1)$$

$$dS_i(t) = S_i(t) \left[ \mu_i(t)dt + \sum_{j=1}^d \sigma_{ij}(t)dW_j(t) \right] \quad ; \quad S_i(0) = s_{i0} > 0, \quad (1.1.2)$$

where

$r(t)$  is the riskless interest rate,

$\mu(t) = (\mu_1(t), \dots, \mu_n(t))'$  denotes the drift vector and

$\sigma(t) = \begin{pmatrix} \sigma_{11}(t) & \dots & \sigma_{1d}(t) \\ \vdots & \ddots & \vdots \\ \sigma_{n1}(t) & \dots & \sigma_{nd}(t) \end{pmatrix}$  denotes the volatility matrix.

Thereby, we make the following standing mathematical assumptions:

- $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$  is a filtered probability space satisfying the usual conditions.<sup>1</sup>
- $\{W_t, \mathcal{F}_t\}_{t \in [0, T]}$  is a (multi-dimensional) Brownian motion of appropriate dimension adapted to the Brownian filtration.
- The coefficients  $r$ ,  $\mu$  and  $\sigma$  are progressively measurable processes being uniformly bounded.

The Black-Scholes model makes following simplifications:

$$n = d = 1 \text{ and } \mu, r, \sigma \text{ constant with } \sigma > 0.$$

This means, that the MMA and the risky asset behave as follows:

$$dB(t) = B(t)r dt \quad ; \quad B(0) = 1 \quad (1.1.3)$$

$$dS(t) = S(t) [\mu dt + \sigma dW(t)] \quad ; \quad S(0) = s_0 > 0. \quad (1.1.4)$$

---

<sup>1</sup>A filtered probability space  $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$  satisfies the usual conditions, if

- (i) it is complete,
- (ii)  $\mathcal{F}_0$  contains all the  $\mathbb{P}$ -null sets of  $\mathcal{F}$ , and
- (iii)  $\{\mathcal{F}_t\}$  is right-continuous.

We now want to price a European call written on a stock  $S$  with strike  $K$ , maturity  $T$  and payoff

$$C(T) = \max \{S(T) - K; 0\}.$$

### Derivation of the Black-Scholes PDE

#### Assumptions:

- No payout is made during the lifetime of the call.
- There exists a  $C^{1,2}$ -function  $f = f(t, s)$ , such that the time- $t$  price of the call is given by  $C(t) = f(t, S(t))$ .

#### First step: Call price dynamics

Ito's formula yields:

$$\begin{aligned} dC(t) &= f_t(t, S(t))dt + f_s(t, S(t))dS(t) + 0.5f_{ss}(t, S(t))d\langle S \rangle_t \\ &= f_t(t, S(t))dt + f_s(t, S(t)) \cdot S(t) [\mu dt + \sigma dW(t)] + 0.5f_{ss}(t, S(t)) \cdot S(t)^2 \sigma^2 dt \\ &= [f_t(t, S(t)) + f_s(t, S(t)) \cdot S(t)\mu + 0.5f_{ss}(t, S(t)) \cdot S(t)^2 \sigma^2] dt + f_s(t, S(t)) \cdot S(t)\sigma dW(t), \end{aligned}$$

since  $dS(t) = S(t) [\mu dt + \sigma dW(t)]$  and  $d\langle S \rangle_t = S(t)\sigma \cdot S(t)\sigma dt = S(t)^2 \sigma^2 dt$ .

In the following we skip the time and stock dependences and write as short hand notation:

$$dC(t) = (f_t + f_s S \mu + 0.5 f_{ss} S^2 \sigma^2) dt + f_s S \sigma dW.$$

#### Second step: Construction of a riskless portfolio

As next we construct a riskless portfolio in order to use the no-arbitrage paradigm for the portfolio. If the portfolio is (locally) risk-free we know that for the wealth of the investor  $X^\varphi(t)$  holds  $dX^\varphi(t) = rX^\varphi(t)$ . Using this we can derive a PDE for the call price. Consider a self-financing trading strategy  $\varphi(t) = (\varphi_B(t), \varphi_S(t), \varphi_C(t))$  and choose  $\varphi_C(t) = -1$ .

Since the trading strategy is self-financing, the wealth of the investor reads:

$$\begin{aligned} dX^\varphi(t) &= \varphi_B(t)dB(t) + \varphi_S(t)dS(t) - dC(t) \\ &= \varphi_B r B dt + \varphi_S S (\mu dt + \sigma dW) - (f_t + f_s S \mu + 0.5 f_{ss} S^2 \sigma^2) dt - f_s S \sigma dW \\ &= \underbrace{(\varphi_B r B + \varphi_S S \mu - f_t - f_s S \mu - 0.5 f_{ss} S^2 \sigma^2)}_{(*)} dt + \underbrace{(\varphi_S S \sigma - f_s S \sigma)}_{(**)} dW. \end{aligned}$$

**Choosing**  $\varphi_S(t) = f_s(t)$  leads to  $(**) = 0$ . Hence, the strategy is locally risk-free, which implies by the no arbitrage paradigm that the wealth of the investor has to behave like the MMA, i.e.  $(*) = X^\varphi(t)r$ .

Rewriting  $(*)$  yields for the drift coefficient:

$$\begin{aligned} (*) &= \varphi_B r B + \varphi_S S \mu - f_t - f_s S \mu - 0.5 f_{ss} S^2 \sigma^2 \\ &= \varphi_B r B - f_t - 0.5 f_{ss} S^2 \sigma^2 \quad (\varphi_S(t) = f_s(t)) \\ &= r(\varphi_B B + \varphi_S S - C) - r f_s S + r C - f_t - 0.5 f_{ss} S^2 \sigma^2 \\ &= r \underbrace{(\varphi_B B + \varphi_S S - C)}_{=X^\varphi} \underbrace{- r f_s S + r f - f_t - 0.5 f_{ss} S^2 \sigma^2}_{(***)}. \end{aligned}$$

Since  $(*)$  has to be equal to  $X^\varphi(t)r$ , it follows that  $(***) = 0$ , i.e.

$$r f_s(t, S(t)) S(t) - r f(t, S(t)) + f_t(t, S(t)) + 0.5 f_{ss}(t, S(t)) S^2(t) \sigma^2 = 0.$$



**Theorem 1.1 (PDE of Black-Scholes (1973)).**

In an arbitrage-free market the price function  $f(t, s)$  of a European call satisfies the Black-Scholes PDE

$$f_t(t, s) + rsf_s(t, s) + 0.5s^2\sigma^2f_{ss}(t, s) - rf(t, s) = 0, \quad (1.1.5)$$

$(t, s) \in [0, T] \times \mathbb{R}_+$ , with terminal condition  $f(T, s) = \max\{s - K, 0\}$ .

**Proof:**

See the calculations above. ■

**Remark.** We have not used the terminal condition to derive the PDE. Therefore, every European (path-independent) option satisfies this PDE. Of course, the terminal condition will be different.

**Closed-form solution for European call**

The above derivation is only valid if there exists a  $C^{1,2}$ -solution of equation (1.1.5). So we should pose us the question: *How can we solve this PDE?*

**1st approach: Merton(1973).**

Solve the PDE by transforming it to the heat equation  $u_t = u_{xx}$ , which has a well known solution. This approach is explicitly done in [Wilmott, 2006] on pp. 102-108.

**2nd approach:**

Apply the *Feynman-Kac representation*, which states that the time- $t$  solution of the Black-Scholes PDE is given by

$$C(t) = \mathbb{E}[\max\{Y(T) - K; 0\} | Y(t) = S(t)] \cdot e^{-r(T-t)},$$

where  $Y$  satisfies the SDE (Q-Dynamics of the asset price)

$$dY(s) = Y(s) [rds + \sigma dW(s)] \quad \forall s \geq t \quad \text{with initial condition} \quad Y(t) = S(t).$$

A proof of the Feynman-Kac representation can be found in [Kraft, 2005] on pp. 52. Applying this representation, we get the well-known formula of Black-Scholes, which was a breakthrough in financial mathematics.

**Theorem 1.2 (Black-Scholes Formula).**

(i) The price of a European call is given by

$$C(t) = C(t, S(t), \sigma, r, K, T) = S(t) \cdot \Phi(d_1(t)) - K \cdot \Phi(d_2(t)) \cdot e^{-r(T-t)}, \quad (1.1.6)$$

where

$$d_1(t) = \frac{\ln\left(\frac{S(t)}{K}\right) + (r + 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}}, \quad d_2(t) = d_1(t) - \sigma\sqrt{T - t} \quad \text{and}$$

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{1}{2}x^2} dx \quad \text{is the cumulative standard normal distribution function.}$$

(ii) The trading strategy  $\varphi := (\varphi_B, \varphi_S)$  with

$$\varphi_B(t) = \frac{C(t) - f_s(t, S(t)) \cdot S(t)}{B(t)}$$

$$\varphi_S(t) = f_s(t, S(t))$$

is a self-financing replication strategy for the call.

**Proof:**

ad(i):

The idea of this proof follows quite along [Korn & Korn, 2001], pp. 102.

Due to the Feynman-Kac representation we have to calculate

$$C(t) = [\max\{Y(T) - K; 0\} | Y(t) = S(t)] \cdot e^{-r(T-t)}, \quad (+)$$

where  $Y$  satisfies the SDE

$$dY(s) = Y(s) [rds + \sigma dW(s)] \quad \forall s \geq t \quad \text{with initial condition} \quad Y(t) = S(t). \quad (++)$$

By variation of constants, the solution of  $(++)$  for  $s = T$  reads:

$$Y(T) = Y(t) \cdot e^{(r-0.5\sigma^2)(T-t) + \sigma(W(T)-W(t))} = S(t) \cdot e^{(r-0.5\sigma^2)(T-t) + \sigma(W(T)-W(t))}.$$

Substituting this into  $(+)$  yields:

$$\begin{aligned} C(t) &= \left[ \max\{Y(t) \cdot e^{(r-0.5\sigma^2)(T-t) + \sigma(W(T)-W(t))} - K; 0\} | Y(t) = S(t) \right] \cdot e^{-r(T-t)} \\ &= \left[ \max\{S(t) \cdot e^{(r-0.5\sigma^2)(T-t) + \sigma(W(T)-W(t))} - K; 0\} \right] \cdot e^{-r(T-t)} \end{aligned}$$

Thereby it holds:

$$\begin{aligned} &S(t) \cdot e^{(r-0.5\sigma^2)(T-t) + \sigma(W(T)-W(t))} - K > 0 \\ \iff &W(T) - W(t) > \frac{1}{\sigma} \left\{ \ln \left( \frac{K}{S(t)} \right) - (r - 0.5\sigma^2)(T - t) \right\} =: \hat{K} \end{aligned}$$

Since  $W(T) - W(t) \sim \mathcal{N}(0, T - t)$  by the definition of Brownian motion, we can write:

$$\begin{aligned} C(t) &= e^{-r(T-t)} \int_{\hat{K}}^{\infty} \left\{ (S(t) \cdot e^{(r-0.5\sigma^2)(T-t) + \sigma x} - K) \cdot \frac{1}{\sqrt{2\pi(T-t)}} \cdot e^{-\frac{x^2}{2(T-t)}} \right\} dx \\ &= S(t) \int_{\hat{K}}^{\infty} \frac{1}{\sqrt{2\pi(T-t)}} e^{-0.5\sigma^2(T-t) + \sigma x - \frac{x^2}{2(T-t)}} dx - K \int_{\hat{K}}^{\infty} \frac{1}{\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{x^2}{2(T-t)}} dx \\ &\stackrel{(\circ)}{=} S(t) \int_{\hat{K}}^{\infty} \frac{1}{\sqrt{2\pi(T-t)}} e^{-\frac{(x - \sigma(T-t))^2}{2(T-t)}} dx - K \cdot e^{-r(T-t)} \int_{\hat{K}}^{\infty} \frac{1}{\sqrt{2\pi(T-t)}} e^{-\frac{x^2}{2(T-t)}} dx \\ &\stackrel{(\circ\circ)}{=} S(t) \underbrace{\left[ 1 - \Phi \left( \frac{\hat{K} - \sigma(T-t)}{\sqrt{T-t}} \right) \right]}_{(*)} - K \cdot e^{-r(T-t)} \underbrace{\left[ 1 - \Phi \left( \frac{\hat{K}}{\sqrt{T-t}} \right) \right]}_{(**)} \end{aligned}$$

Note that  $(\circ)$  holds due to

$$-0.5\sigma^2(T-t) + \sigma x - \frac{x^2}{2(T-t)} = -\frac{1}{2(T-t)} [x^2 - 2\sigma(T-t)x + \sigma^2(T-t)^2] = -\frac{[x - \sigma(T-t)]^2}{2(T-t)}$$

and  $(\circ\circ)$  holds because if  $\varphi(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^y e^{\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} dx$  is the cumulative normal distribution function with expectation  $\mu$  and variance  $\sigma^2$  and  $\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{1}{2}x^2} dx$  is the cumulative standard normal distribution function, then it holds:

$$\varphi(y) = \Phi \left( \frac{y - \mu}{\sigma} \right).$$

We finally calculate (\*) and (\*\*):

$$\begin{aligned}
(*) &= 1 - \Phi \left( \frac{\frac{1}{\sigma} \left\{ \ln \left( \frac{K}{S(t)} \right) - (r - 0.5\sigma^2)(T - t) \right\} - \sigma(T - t)}{\sqrt{T - t}} \right) \\
&= 1 - \Phi \left( \frac{\ln \left( \frac{K}{S(t)} \right) - (r - 0.5\sigma^2)(T - t) - \sigma^2(T - t)}{\sigma\sqrt{T - t}} \right) = 1 - \Phi \left( \frac{\ln \left( \frac{K}{S(t)} \right) - (r + 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}} \right) \\
&= \Phi \left( \frac{\ln \left( \frac{S(t)}{K} \right) + (r + 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}} \right) = \Phi(d_1(t)) \\
\\ 
(**) &= 1 - \Phi \left( \frac{\frac{1}{\sigma} \left\{ \ln \left( \frac{K}{S(t)} \right) - (r - 0.5\sigma^2)(T - t) \right\}}{\sqrt{T - t}} \right) = 1 - \Phi \left( \frac{\ln \left( \frac{K}{S(t)} \right) - (r - 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}} \right) \\
&= \Phi \left( \frac{\ln \left( \frac{S(t)}{K} \right) + (r - 0.5\sigma^2)(T - t)}{\sigma\sqrt{T - t}} \right) = \Phi(d_2(t)) = \Phi(d_1(t) - \sigma\sqrt{T - t})
\end{aligned}$$

This completes the proof of (i).

ad(ii):

We have

$$X^\varphi(t) = \varphi_B(t)B(t) + \varphi_S(t)S(t) = C(t) \quad . \quad (\diamond)$$

Hence,  $(\varphi_B, \varphi_S)$  replicates the call and is self-financing because

$$\begin{aligned}
dX^\varphi &\stackrel{(\diamond)}{=} dC \stackrel{\text{Ito}}{=} (f_t + f_s S \mu + 0.5 f_{ss} S^2 \sigma^2) dt + f_s S \sigma dW \\
&\stackrel{(\diamond\diamond)}{=} (r[C - f_s S] + f_s S \mu) dt + f_s S \sigma dW \\
&\stackrel{\text{Def.}}{=}_{\varphi_B, \varphi_S} (rB\varphi_B + \varphi_S S \mu) dt + \varphi_S S \sigma dW \\
&= (\varphi_B r B dt + \varphi_S S [\mu dt + S \sigma dW]) = \varphi_B dB + \varphi_S dS .
\end{aligned}$$

$(\diamond\diamond)$  holds due to the Black-Scholes PDE:

$$f_t + 0.5 f_{ss} S^2 \sigma^2 = r[C - f_s S] .$$

This completes the proof of (ii). ■

#### Remarks.

- The strategy in (ii) is said to be a **delta hedging strategy**.
- If the EMM  $\mathbb{Q}$  is unique, i.e.  $n = d$  (“# stocks = # sources of risk”) and  $\sigma$  is uniformly positive definite (which is covered by the assumptions  $n = d = 1$  and  $\sigma > 0$  of Black and Scholes), the delta hedging strategy is the **unique** replication strategy.

## 1.2 Some stylized Facts and implied Volatility

An erroneous assumptions of the Black-Scholes model is that the volatility of the underlying is constant. One has only to think of the stock market crash of October 1987. Also statistical tests strongly reject the assumption that a constant volatility process could have generated stock market returns.

Moreover, some so called stylized facts and the fact that options with different strikes and maturities have different Black-Scholes implied volatilities point out that the volatility is non-constant.

In this subsection, these stylized facts and implied volatility will be explained.

### Some stylized Facts

Stock prices, exchange and interest rates and other financial time series have some empirical substantiated properties which distinguish them from other time series. These empirical properties are called **stylized facts**.

In the following, we will only investigate stylized facts, which give evidence that the volatility is non-constant. Of course, there are many more stylized facts than mentioned.

To introduce the term of a stylized fact, we first need some basic definitions:

#### Definition 1.3:

A **time series (in discrete time)** is defined as

- a sorted sequence of observations observed at discrete time points.
- a realization of a stochastic process  $\{X(n)\}_{n \in \mathbb{N}}$ .

#### Definition 1.4:

Let  $\{S(n)\}_{n \in \mathbb{N}}$  be the time series of the asset price and assume that no dividends are paid.

(i) The (discrete) **return** from time  $n-1$  to time  $n$  is defined (for  $n \geq 1$ ) as

$$R(n) := \frac{S(n) - S(n-1)}{S(n-1)} \quad (1.2.1)$$

(ii) The (discrete) **log return** from time  $n-1$  to time  $n$  is defined (for  $n \geq 1$ ) as

$$r(n) := \ln \frac{S(n)}{S(n-1)} \quad (1.2.2)$$

**Remark.** For small price changes, the log returns are a good approximation for the returns, since the following holds by using the definition of returns and a Taylor expansion at the expansion point  $R(n) = 0$ :

$$\begin{aligned} r(n) &= \ln \frac{S(n)}{S(n-1)} = \ln(1 + R(n)) \\ &= \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \cdot R^l(n) \quad \text{for } |R(n)| < 1 \\ &= R(n) - \frac{R^2(n)}{2} + \frac{R^3(n)}{3} \mp \dots \\ &\approx R(n) \quad , \quad \text{if } |R(n)| \text{ small enough (which means small price changes).} \end{aligned}$$

This is often the case if one examines financial time series with a high frequency, i.e. for example daily data.

***Stylized fact 1: Volatility Clustering***

In Figure 1, the daily log returns  $\{r(n)\}_{n \geq 1}$  of the DAX Index from 1990-1992 are plotted against time. One can see, that large moves follow large moves and small moves follow small moves. This is the so called volatility clustering.

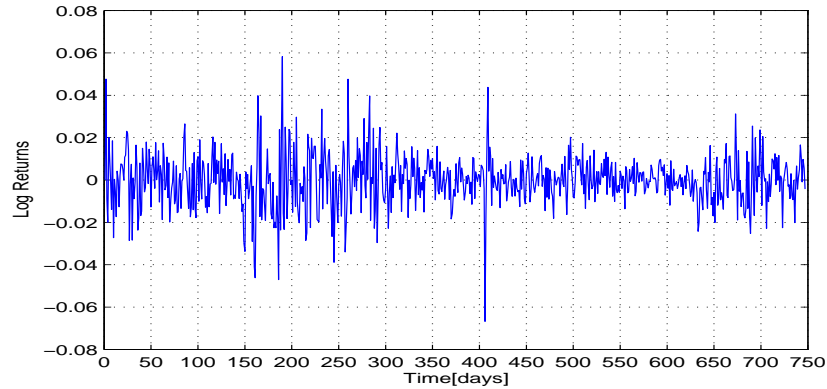


Figure 1: DAX daily Log Returns from 1990-1992

***Stylized fact 2: Leverage Effect***

Empirical data show that for in particular for asset returns negative news in terms of bills or large losses have a stronger influence on the perception of current risk and so the volatility than positive news in terms of large gains. This means, that the volatility responds asymmetrically to the sign of shocks. This behaviour is referred to as leverage effect.

This fact was first mentioned by Fisher Black in 1976, who made the following statement:

*“A drop in the value of the firm will cause a negative return on its stock, and will usually increase the leverage of the stock. [...] That rise in the debt-equity ratio will surely mean a rise in the volatility of the stock.”*

*Fisher Black (1976)*

***Stylized fact 3: Heavy Tails***

In Figure 2, we plot the frequency distribution of the DAX daily log returns  $\{r(n)\}_{n \geq 1}$  from 1990-1992 and compare it with the normal distribution. The plot shows that the frequency distribution is highly peaked and fat tailed.

The normplot of the DAX daily log returns in Figure 3 just shows how extreme the tails of the empirical distribution are compared to the normal distribution.

For explanation:

The purpose of a normplot (= normal probability plot) is to graphically assess whether the data could come from a normal distribution. If the DAX returns were normal distributed, the plot would be a straight line (just like the dotted red line in Figure 3).

Heavy tails can also be described in a more mathematical way. To do this, we need some definitions.

**Definition 1.5:**

Let  $X$  be a (real-valued) random variable with  $[X^4] < \infty$ . We then define

- the **kurtosis** of  $X$  as

$$\kappa(X) := \frac{[(X - [X])^4]}{(\text{Var}[X])^2} \quad (1.2.3)$$

- the **excess** of  $X$  as

$$\varepsilon(X) := \kappa(X) - 3 \quad (1.2.4)$$

**Remark.** For discrete data  $\{X(n)\}_{n=1,\dots,N}$  (like returns or log returns), the kurtosis is estimated by the *sample kurtosis* which is defined as follows:

$$\hat{\kappa}_N = \frac{1}{(\hat{\text{Var}}_N)^2} \cdot \frac{1}{N} \sum_{n=1}^N (X(n) - \bar{\mu}_N)^4, \quad (1.2.5)$$

where

- $\bar{\mu}_N = \frac{1}{N} \sum_{n=1}^N X(n)$  = sample mean
- $\hat{\text{Var}}_N = \frac{1}{N} \sum_{n=1}^N (X(n) - \bar{\mu}_N)^2$  = sample variance .

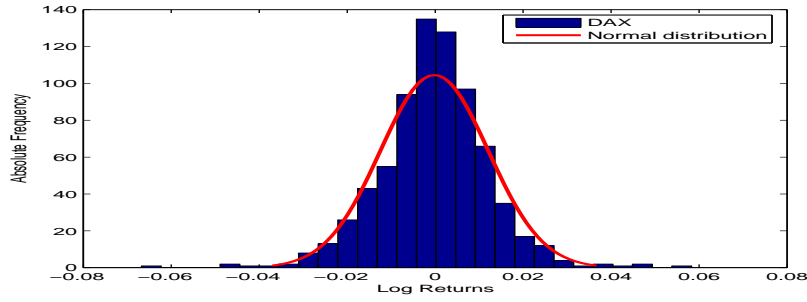


Figure 2: Frequency Distribution of DAX daily Log Returns from 1990-1992

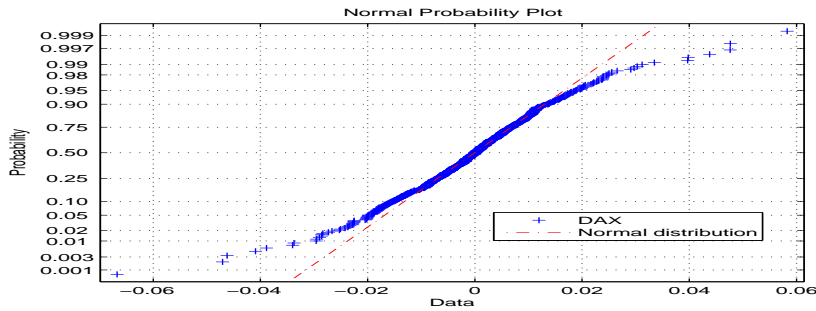


Figure 3: Normplot of DAX daily Log Returns

The following proposition gives a special property of the kurtosis of normal distributed random variables.

**Proposition 1.6 (Kurtosis of normal distributed Random Variables).**

Let  $X$  be a (real-valued) random variable. If  $X \sim \mathcal{N}(\mu, \sigma^2)$ , then it holds:

- (i)  $\kappa(X) = 3$
- (ii)  $\varepsilon(X) = 0$ .

**Proof:**

ad(i):

$X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow [X] = \mu ; \text{Var}[X] = \sigma^2 ; p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ , where  $p$  is the density of the normal distribution.

By definition, we have:

$$\begin{aligned}
 \kappa(X) &= \frac{[(X - [X])^4]}{(\text{Var}[X])^2} = \frac{[(X - \mu)^4]}{\sigma^4} \\
 &= \frac{\int_{-\infty}^{+\infty} (x - \mu)^4 \cdot p(x) \, dx}{\sigma^4} \\
 &\downarrow p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
 &= \frac{\int_{-\infty}^{+\infty} (x - \mu)^4 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx}{\sigma^4} \\
 &= \frac{1}{\sigma^5 \sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} (x - \mu)^4 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \\
 &\downarrow \text{The integrand is an even function and therefore symmetric to the y-axis.} \\
 &= \frac{2}{\sigma^5 \sqrt{2\pi}} \cdot \int_0^{+\infty} (x - \mu)^4 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \\
 &\downarrow \text{Substitution: } y := x - \mu \text{ resp. } x := y + \mu \Rightarrow \frac{dx}{dy} = 1 \Rightarrow dx = dy \\
 &= \frac{2}{\sigma^5 \sqrt{2\pi}} \cdot \int_0^{+\infty} y^4 \cdot e^{-\frac{y^2}{2\sigma^2}} \, dy \\
 &\downarrow \text{Due to [Bronstein et al., 2000], p.1084 it holds:}^2 \\
 &\quad \int_0^{\infty} y^4 \cdot e^{-\frac{y^2}{2\sigma^2}} \, dy = \frac{\Gamma(\frac{5}{2})}{2 \cdot (\frac{1}{2\sigma^2})^{\frac{5}{2}}} = (\sqrt{2})^3 \cdot \sigma^5 \cdot \Gamma\left(\frac{5}{2}\right) \stackrel{\Gamma(\frac{5}{2})=\frac{3}{4}\sqrt{\pi}}{=} \frac{\sigma^5 \sqrt{2\pi}}{2} \cdot 3 \\
 &= \frac{2}{\sigma^5 \sqrt{2\pi}} \cdot \frac{\sigma^5 \sqrt{2\pi}}{2} \cdot 3 = 3.
 \end{aligned}$$

ad(ii):

$$\varepsilon(X) = \kappa(X) - 3 \stackrel{(i)}{=} 3 - 3 = 0. \blacksquare$$

**Definition 1.7:**

Let  $X$  be a (real-valued) random variable. If  $\kappa(X) > 3$ , then its distribution is said to be **leptokurtic**.

**Remark.** The kurtosis of financial data is often bigger than 3 and so the (frequency) distribution is leptokurtic. This indicates, that values close to zero and in particular very large positive

<sup>2</sup>Thereby  $\Gamma$  denotes the gamma function, which is defined as follows:  $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt \quad (x > 0)$ .

(or negative) values occur with a (compared to the normal distribution) higher probability. This connects the fat tails with the kurtosis. Leptokurticity is a measure for fat tails and central peaks. **The higher the kurtosis the heavier are the tails** ( $\longrightarrow$  more extreme events).

Table 1 shows the sample kurtosis of several different shares traded at the DAX and the DAX itself from 1990-1992. One can observe that the kurtosis is clearly bigger than 3 in every case, which is empirical evidence for the heavy tails of financial data.

	DAX	Deutsche Bank	Dresdner Bank	BASF	Bayer	BMW	VW
$\hat{\kappa}_{N=746}$	5.8353	7.0517	8.7896	7.6028	6.3665	6.7934	7.1756

Table 1: Sample Kurtosis of DAX Shares and DAX

*Heavy tails and a high central peak are characteristics of mixtures of distributions with different variances (and so volatilities). So this fact implies a non-constant volatility, too.*

#### **Stylized fact 4: Skewness**

Financial data also show an asymmetry in their probability distribution. This property is measured by the skewness of a random variable.

##### **Definition 1.8:**

Let  $X$  be a (real-valued) random variable with  $[X^3] < \infty$ . We then define the **skewness** of  $X$  as

$$\gamma(X) := \frac{[(X - [X])^3]}{(\text{Var}[X])^{\frac{3}{2}}} . \quad (1.2.6)$$

**Remark.** Analogously to the kurtosis, for discrete data  $\{X(n)\}_{n=1, \dots, N}$ , the skewness is estimated by the *sample skewness* which is defined as follows:

$$\hat{\gamma}_N = \frac{1}{(\hat{\text{Var}}_N)^{\frac{3}{2}}} \cdot \frac{1}{N} \sum_{n=1}^N (X(n) - \bar{\mu}_N)^3 , \quad (1.2.7)$$

where  $\bar{\mu}_N$  = sample mean and  $\hat{\text{Var}}_N$  = sample variance.

The following proposition gives a property of the skewness of normal distributed random variables.

##### **Proposition 1.9 (Skewness of normal distributed Random Variables).**

Let  $X$  be a (real-valued) random variable. If  $X \sim \mathcal{N}(\mu, \sigma^2)$ , then it holds:

$$\gamma(X) = 0 .$$



**Proof:**

Analogously to proposition 1.6 we can conclude:

$$\begin{aligned}
\gamma(X) &= \frac{[(X - [X])^3]}{(\text{Var}[X])^{\frac{3}{2}}} = \frac{[(X - \mu)^3]}{\sigma^3} = \frac{\int_{-\infty}^{+\infty} (x - \mu)^3 \cdot p(x) dx}{\sigma^3} \\
&= \frac{\int_{-\infty}^{+\infty} (x - \mu)^3 \cdot \frac{1}{\sqrt{2\pi}\sigma^2} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}{\sigma^3} = \frac{1}{\sigma^4\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} (x - \mu)^3 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&\downarrow \text{ Substitution: } y := x - \mu \text{ resp. } x := y + \mu \Rightarrow \frac{dx}{dy} = 1 \Rightarrow dx = dy \\
&= \frac{1}{\sigma^4\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} y^3 \cdot e^{-\frac{y^2}{2\sigma^2}} dy = \frac{1}{\sigma^4\sqrt{2\pi}} \cdot \underbrace{\int_{-\infty}^{+\infty} ye^{-\frac{y^2}{2\sigma^2}} \cdot y^2 dy}_{(*)}
\end{aligned}$$

The integral (\*) can now easily be solved by integrating by parts:

$$\int_{-\infty}^{+\infty} ye^{-\frac{y^2}{2\sigma^2}} \cdot y^2 dy = \underbrace{\left[ -\sigma^2 e^{-\frac{y^2}{2\sigma^2}} \cdot y^2 \right]_{-\infty}^{+\infty}}_{=0} - \int_{-\infty}^{+\infty} -\sigma^2 e^{-\frac{y^2}{2\sigma^2}} \cdot 2y dy = \left[ -2\sigma^4 e^{-\frac{y^2}{2\sigma^2}} \right]_{-\infty}^{+\infty} = 0$$

This completes the proof. ■

**Remark.** The more the skewness differs from zero the more asymmetric is the (frequency) distribution of the random variable or the time series data. If  $\gamma(X) < 0$  (skewed left) the left tail is long relative to the right tail and if  $\gamma(X) > 0$  (skewed right) the right tail is long relative to the left tail.

Table 2 shows the sample skewness of different DAX shares and the DAX index. Any value is obviously different from zero, which points out that financial data are in fact skewed.

	DAX	Deutsche Bank	Dresdner Bank	BASF	Bayer	BMW	VW
$\hat{\gamma}_{N=746}$	-0.0587	-0.1577	-0.7009	-0.2363	-0.1632	-0.1480	-0.0584

Table 2: Sample Skewness of DAX Shares and DAX

*The fact that the skewness different is from zero for all shares listed above is an indication for the non-normality of financial data. This shows that the Black-Scholes model is mis-specified since it assumes normality of the returns.*

The stylized facts 1-4 explained above strongly indicate that the volatility cannot be constant from a statistical point of view and that the Black-Scholes model cannot reflect all empirical observations. It is rather advisable to model volatility even as a random variable.

We get further evidence for the fact that the volatility is non-constant when we look at the volatilities which are given by the market.

**Implied Volatility**

The concept of (Black-Scholes) implied volatility can be defined according to market prices

or according to option prices that come out of a more advanced equity price model after calibrating it to market prices. In the following we explain the concept of implied volatilities that are determined from market prices of options.

From theorem 1.2 we know, that the price of a European call in the Black-Scholes framework of constant volatility simply is

$$C(t) = C(t, S(t), \sigma, r, K, T) = S(t) \cdot \Phi(d_1(t)) - K \cdot \Phi(d_2(t)) \cdot e^{-r(T-t)},$$

where

$$d_1(t) = \frac{\ln\left(\frac{S(t)}{K}\right) + (r + 0.5\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2(t) = d_1(t) - \sigma\sqrt{T-t}.$$

We have given the function  $C$  six arguments. All of them except of the volatility are directly observable. If we know  $\sigma$ , we can calculate the option price. Conversely, if we know the option price  $C$ , then we can calculate  $\sigma$ . Since this is the case, we will exploit the relationship between prices and volatility to determine the volatility from the *market prices*.

Let  $C_M(K_i, T_j)$  denote the market price of a call with strike  $K_i$  and maturity  $T_j$ . In a liquid market, we can observe a set of market prices  $(C_M(K_i, T_j))_{i=1, \dots, N; j=1, \dots, M}$  with strikes  $(K_i)_{i=1, \dots, N}$  and maturities  $(T_j)_{j=1, \dots, M}$ .

**Assumption:** *The market “knows” the right volatility.*

If we now require  $C_M$  to equal  $C$ , we can calculate the volatility from the market prices:

$$C(t, S(t), \sigma_{\text{imp}}, r, K_i, T_j) \stackrel{!}{=} C_M(K_i, T_j) \quad \text{for } i = 1, \dots, N \text{ and } j = 1, \dots, M \quad . \quad (1.2.8)$$

$\sigma_{\text{imp}}$  is the value of  $\sigma$  for which the theoretical Black-Scholes option value matches the market price of the option and is only implicitly determined by equation (1.2.8). It has to be computed numerically since it is impossible to resolve equation (1.2.8) to  $\sigma$ . This is called (Black-Scholes) **implied volatility**.

Since we have a set of market prices  $(C_M(K_i, T_j))_{i=1, \dots, N; j=1, \dots, M}$  for different strikes  $(K_i)_{i=1, \dots, N}$  and maturities  $(T_j)_{j=1, \dots, M}$ ,  $\sigma_{\text{imp}}$  obviously depends on strike and maturity:

$$\sigma_{\text{imp}} = \sigma_{\text{imp}}(K_i, T_j) \quad \text{for } i = 1, \dots, N \text{ and } j = 1, \dots, M \quad .$$

**Definition 1.10:**

A plot of the implied volatility  $(\sigma_{\text{imp}}(K_i, T_j))_{i=1, \dots, N; j=1, \dots, M}$  of an option or an index of options as a function of its strike price  $(K_i)_{i=1, \dots, N}$  and its maturity  $(T_j)_{j=1, \dots, M}$  is called a **volatility surface**.

Figure 4 shows the volatility surface of the S&P 500 Index and the N225 Index, both on July 12, 2002. One can see, that *options with different strikes and maturities have different implied volatilities*. Figure 5 shows in contrast the volatility surface for a constant Black-Scholes volatility. Comparing the two figures, we can see that the term of implied volatility also clearly gives evidence for the fact that the assumption of a constant volatility  $\sigma$  in the Black-Scholes model is wrong.

Finally, we can conclude that our examination of stylized facts and implied volatility has shown that one should not model volatility as constant. This is the reason why more sophisticated

equity price models have been developed; a variety of them from local volatility models over stochastic volatility models (especially Heston model) to the Bergomi model will be investigated in this diploma thesis.

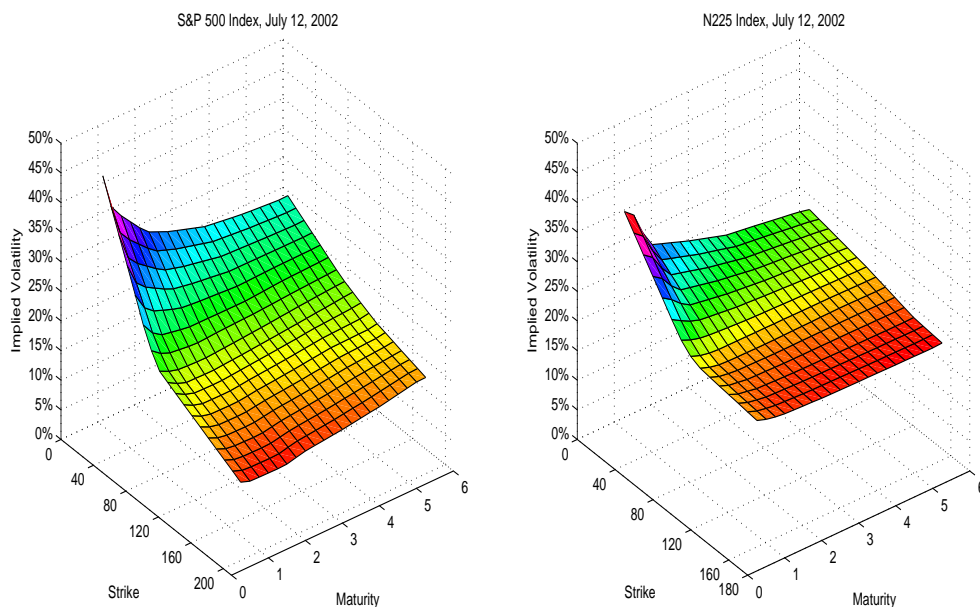


Figure 4: Volatility Surface of S&P 500 Index and N225 Index July 12, 2002

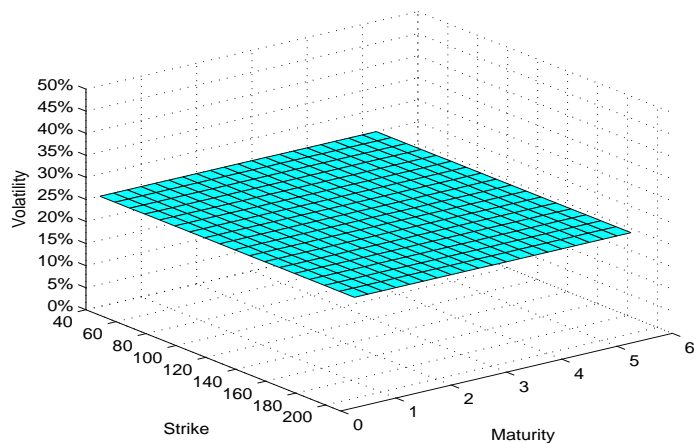


Figure 5: Volatility Surface assumed by Black and Scholes

### Digression 1 : Volatility Smiles, Skews and Frowns

We will give here an exact definition of a volatility smile, a volatility skew and a volatility frown so that they can be clearly distinguished during the course of this diploma thesis.

In section 1.2 we defined a volatility surface. Now we fix the maturity  $T$  and allow the market price to depend only on its strike price  $K$ . We then denote the market price of a call with fixed maturity and strike price  $K_i$  with  $C_M^T(K_i)$  and analogously the set of market prices with fixed maturity  $T$  and varying strike prices  $(K_i)_{i=1,\dots,N}$  with  $(C_M^T(K_i))_{i=1,\dots,N}$ . Thus, the implied volatility  $\sigma_{\text{imp}}$  depends for fixed maturity only on  $(K_i)_{i=1,\dots,N}$ :

$$\sigma_{\text{imp}} = \sigma_{\text{imp}}^T(K_i) \quad \text{for } i = 1, \dots, N \quad .$$

#### Definition 1.11:

A plot of the implied volatility  $(\sigma_{\text{imp}}^T(K_i))_{i=1,\dots,N}$  for fixed maturity  $T$  of an option or an index of options as a function of its strike price  $(K_i)_{i=1,\dots,N}$  is in general called a **volatility smile**.

There are different shapes of the volatility smile for different markets. We distinguish here three different shapes:

#### Foreign Currency Markets:

The general shape of the volatility smile that originates from the implied volatilities of foreign currency options is shown in Figure 6.

The implied volatility is relatively low for at-the-money options and it is symmetric around  $K = S$  because it becomes progressively higher as an option moves either into the money or out of the money. This shape is called **volatility smile**.



Figure 6: Scheme of a Volatility Smile

#### Equity Markets:

The shape of the volatility smile that originates from the implied volatilities of equity options is not as symmetric as the shape of the smile of foreign currency options.

The implied volatility decreases as the strike price increases. The implied volatility of low-strike-price options is significantly higher than that of high-strike-price options. This is referred to as a **volatility skew**.

Figure 7 shows the volatility skew of the S & P 500 and the N225 Index for fixed  $T = 0.25$ .

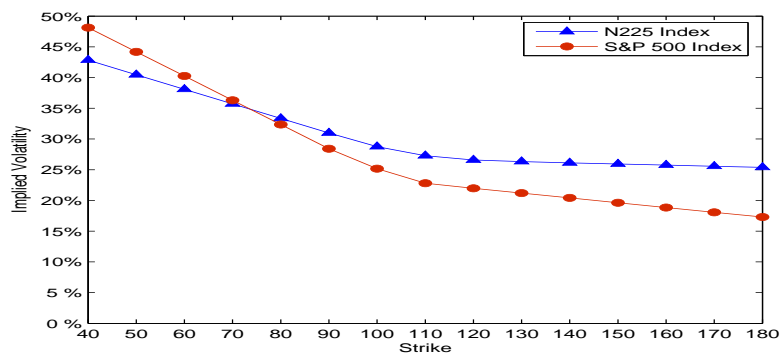


Figure 7: Volatility Skew of S&P 500 Index and N225 Index July 12, 2002 for  $T=0.25$

### Commodity Markets:

The general shape of the volatility smile that originates from the implied volatilities of commodity options is shown in Figure 8. It is upside down to a volatility skew (implied volatility increases as the strike price increases) and is denoted as a **volatility frown**.

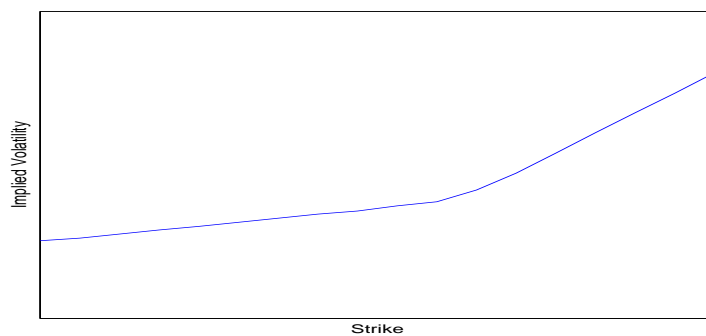


Figure 8: Scheme of a Volatility Frown

*(End of Digression 1)*

## 2 Local Volatility Models

Local volatility models are the easiest models that can produce a smile or a skew. They are still a one factor model and are called **second generation models**.

The basic idea of local volatility models is the following:

Assume that the stock price evolves under the martingale measure according to the risk-neutral diffusion process

$$dS(t) = S(t) [r dt + \sigma(S(t), t) dW^Q(t)] \quad ; \quad S(0) = s_0 > 0 \quad (2.0.1)$$

with a function  $\sigma(S(t), t)$  such that today's option prices can be fitted.

This means, that we still have a 1-D diffusion process that can fit the smile, which is a big advantage of local volatility models.

### 2.1 The Formula of Dupire

In this subsection we derive the formula of Dupire which states how to choose  $\sigma(S(t), t)$  when today's market prices of European call options are given for all possible strikes and maturities such that the smile can be fitted.

#### Assumptions:

- We have a continuous set of market prices of European call options of all strikes and maturities and denote it with  $(C(K, T))_{K, T \in \cdot}$ .
- The stock price evolves according to equation (2.0.1) and no dividends are paid.

#### Notation and Preliminaries:

- The asset price and time are now fixed at today's values  $S(t)$  and  $t$ .
- To analyze the probabilistic properties of an SDE of the general form

$$dy = A(y, t)dt + B(y, t)dW(t) \quad , \quad (2.1.1)$$

we introduce the **transition probability density function**  $\varphi(y, t; y', t')$  which is defined by

$$P(a < y < b \text{ at time } t' | y \text{ at time } t) = \int_a^b \varphi(y, t; y', t') dy' \quad \text{with } t' > t. \quad (2.1.2)$$

- We rely in the following on the transition probability density function  $\varphi(S(t), t; S(T), T)$  for the SDE (2.0.1) assumed for the asset price with forward time variable  $t' = T$ .
- During the derivation of the Dupire equation we set  $z := S(t)$  and  $Z := S(T)$  for the sake of readability and therefore the transition probability density function reads  $\varphi(z, t; Z, T)$ .

Since the value of an option is the discounted value of the expected payoff, we can write for a European call:

$$\begin{aligned} C(K, T) &= e^{-r(T-t)} \int_0^\infty \max\{Z - K; 0\} \varphi(z, t; Z, T) dZ \\ &= e^{-r(T-t)} \int_K^\infty (Z - K) \varphi(z, t; Z, T) dZ \end{aligned}$$

Differentiating w.r.t.  $K$  yields:

$$\frac{\partial C}{\partial K} = \frac{\partial C(K, T)}{\partial K} = -e^{-r(T-t)} \int_K^\infty \varphi(z, t; Z, T) dZ$$

Another differentiation w.r.t.  $K$  gives:

$$\frac{\partial^2 C}{\partial K^2} = -e^{-r(T-t)} \cdot (-\varphi(z, t; Z, T)|_{Z=K}) = -e^{-r(T-t)} \cdot \varphi(z, t; K, T)$$

So we get for the transition probability density function:

$$\varphi(z, t; K, T) = e^{r(T-t)} \frac{\partial^2 C}{\partial K^2} \quad . \quad (2.1.3)$$

The transition probability density function evolves according to the Fokker-Planck or Forward Kolmogorov equation, which is a parabolic partial differential equation requiring initial conditions at time  $t$  and to be solved for  $t' > t$ . It describes the transition probability density function  $\varphi(y, t; y', t')$  of the according process attaining the state  $y'$  at time  $t'$  conditional on an initial state  $y$  at time  $t$ .

**Theorem 2.1 (Fokker-Planck Equation).**

Let  $\varphi(y, t; y', t')$  be the transition probability density function defined by equation (2.1.2) corresponding to the general SDE (2.1.1).

Further let  $f \in C^2$  be an arbitrary function of  $y'$  such that

$$\left[ \int_0^T \left( B(y', t') \frac{\partial f(y')}{\partial y'} \right)^2 dt' \right] < \infty,$$

where  $y'$  fullfills the SDE (2.1.1) and assume that

- $\lim_{y' \rightarrow \pm\infty} f(y') A(y', t') \varphi(y, t; y', t') = 0$
- $\lim_{y' \rightarrow \pm\infty} \frac{\partial f(y')}{\partial y'} B(y', t')^2 \varphi(y, t; y', t') = 0$
- $\lim_{y' \rightarrow \pm\infty} f(y') \frac{\partial}{\partial y'} (B(y', t')^2 \varphi(y, t; y', t')) = 0.$

Then  $\varphi(y, t; y', t')$  fullfills the so called **Fokker-Planck equation**:

$$\frac{\partial \varphi(y, t; y', t')}{\partial t'} = \frac{1}{2} \frac{\partial^2}{\partial y'^2} (B(y', t')^2 \varphi(y, t; y', t')) - \frac{\partial}{\partial y'} (A(y', t') \varphi(y, t; y', t')) \quad . \quad (2.1.4)$$

**Proof:**

Since by assumption  $f$  is an arbitrary function of  $y'$ , where  $y'$  fullfills (2.1.1), we get by Ito's formula:

$$\begin{aligned} df(y') &= \frac{\partial f(y')}{\partial y'} dy' + \frac{1}{2} \frac{\partial^2 f(y')}{\partial y'^2} d\langle y' \rangle_{t'} \\ &= \frac{\partial f(y')}{\partial y'} [A(y', t') dt' + B(y', t') dW(t')] + \frac{1}{2} \frac{\partial^2 f(y')}{\partial y'^2} B(y', t')^2 dt' \\ &= \left[ \frac{\partial f(y')}{\partial y'} A(y', t') + \frac{1}{2} \frac{\partial^2 f(y')}{\partial y'^2} B(y', t')^2 \right] dt' + \frac{\partial f(y')}{\partial y'} B(y', t') dW(t') \end{aligned} \quad (*)$$

Taking the expectation of equation (\*) w.r.t.  $\varphi(y, t; y', t')$  on both sides then yields:

$$\begin{aligned} [df(y')] &= \left[ A(y', t') \frac{\partial f(y')}{\partial t'} dt' \right] + \frac{1}{2} \left[ B(y', t')^2 \frac{\partial^2 f(y')}{\partial y'^2} dt' \right] + \left[ B(y', t') \frac{\partial f(y')}{\partial y'} dW(t') \right] \\ &\stackrel{(+)}{=} \left[ A(y', t') \frac{\partial f(y')}{\partial t'} dt' \right] + \frac{1}{2} \left[ B(y', t')^2 \frac{\partial^2 f(y')}{\partial y'^2} dt' \right] \end{aligned} \quad (**)$$

Thereby (+) holds because we have that  $\left[ \int_0^T f(W(t), t) dW(t) \right] = 0$  when  $f$  is in expectation square-integrable, i.e.  $\left[ \int_0^T f^2(W(s), s) ds \right] < \infty$ . The square-integrability is ensured by the assumption of the theorem.

Taking the derivative of (\*\*) w.r.t.  $t'$ , interchanging integration and differentiation and by the definition of the expectation we get:

$$\begin{aligned} \int_{-\infty}^{+\infty} f(y') \frac{d}{dt'} \varphi(y, t; y', t') dy' \\ = \underbrace{\int_{-\infty}^{+\infty} \frac{\partial f(y')}{\partial y'} A(y', t') \varphi(y, t; y', t') dy'}_{=(\diamond)} + \frac{1}{2} \underbrace{\int_{-\infty}^{+\infty} \frac{\partial^2 f(y')}{\partial y'^2} B(y', t')^2 \varphi(y, t; y', t') dy'}_{=(\diamond\diamond)} \end{aligned}$$

Integrating  $(\diamond)$  once by parts yields:

$$\begin{aligned} (\diamond) &= \left[ f(y') A(y', t') \varphi(y, t; y', t') \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f(y') \frac{\partial}{\partial y'} (A(y', t') \varphi(y, t; y', t')) dy' \\ &= - \int_{-\infty}^{+\infty} f(y') \frac{\partial}{\partial y'} (A(y', t') \varphi(y, t; y', t')) dy' \quad \text{by assumption} \end{aligned}$$

Integrating  $(\diamond\diamond)$  twice by parts yields:

$$\begin{aligned} (\diamond\diamond) &= \left[ \frac{\partial f(y')}{\partial y'} B(y', t')^2 \varphi(y, t; y', t') \right]_{-\infty}^{+\infty} + \left[ f(y') \frac{\partial}{\partial y'} (B(y', t')^2 \varphi(y, t; y', t')) \right]_{-\infty}^{+\infty} \\ &\quad + \int_{-\infty}^{+\infty} f(y') \frac{\partial^2}{\partial y'^2} (B(y', t')^2 \varphi(y, t; y', t')) dy' \\ &= \int_{-\infty}^{+\infty} f(y') \frac{\partial^2}{\partial y'^2} (B(y', t')^2 \varphi(y, t; y', t')) dy' \quad \text{by assumption} \end{aligned}$$

So we get:

$$\begin{aligned} \int_{-\infty}^{+\infty} f(y') \frac{d}{dt'} \varphi(y, t; y', t') dy' \\ = - \int_{-\infty}^{+\infty} f(y') \frac{\partial}{\partial y'} (A(y', t') \varphi(y, t; y', t')) dy' + \frac{1}{2} \int_{-\infty}^{+\infty} f(y') \frac{\partial^2}{\partial y'^2} (B(y', t')^2 \varphi(y, t; y', t')) dy' \end{aligned}$$

Since  $f(y')$  was arbitrary, we can finally conclude:

$$\frac{\partial \varphi(y, t; y', t')}{\partial t'} = - \frac{\partial}{\partial y'} (A(y', t') \varphi(y, t; y', t')) + \frac{1}{2} \frac{\partial^2}{\partial y'^2} (B(y', t')^2 \varphi(y, t; y', t')) .$$

This completes the proof. ■



**Remark.** During the derivation of the Dupire equation we skip the dependences of the transition probability density function and write as short hand notation for the Fokker-Planck equation:

$$\frac{\partial \varphi}{\partial t'} = \frac{1}{2} \frac{\partial^2}{\partial y'^2} (B(y', t')^2 \varphi) - \frac{\partial}{\partial y'} (A(y', t') \varphi) . \quad (2.1.5)$$

This means that in our special case  $A(y, t) = S(t)r = zr$  as well as  $B(y, t) = S(t)\sigma(S(t), t) = z\sigma(z, t)$  and  $t' = T$  as well as  $y' = S(T) = Z$ , the Fokker-Planck equation has the following particular form:

$$\frac{\partial \varphi}{\partial T} = \frac{1}{2} \frac{\partial^2}{\partial Z^2} (\sigma(Z, T)^2 Z^2 \varphi) - \frac{\partial}{\partial Z} (rZ \varphi) \quad (2.1.6)$$

Note that at this point  $\sigma$  is still an unknown function of  $S$  and  $t$  and is evaluated at  $t = T$  in the Fokker-Planck equation.

Taking the derivative of  $C(K, T)$  w.r.t.  $T$  yields:

$$\begin{aligned} \frac{\partial C}{\partial T} &= -re^{-r(T-t)} \int_K^\infty (Z - K) \varphi dZ + e^{-r(T-t)} \int_K^\infty \left\{ \frac{\partial}{\partial T} \varphi \right\} (Z - K) dZ \\ &= -rC + e^{-r(T-t)} \int_K^\infty \left\{ \frac{\partial \varphi}{\partial T} \right\} (Z - K) dZ \end{aligned}$$

Using equation (2.1.6) this can be written as:

$$\frac{\partial C}{\partial T} = -rC + e^{-r(T-t)} \int_K^\infty \left\{ \frac{1}{2} \frac{\partial^2}{\partial Z^2} (\sigma(Z, T)^2 Z^2 \varphi) - \frac{\partial}{\partial Z} (rZ \varphi) \right\} (Z - K) dZ$$

By integrating this twice by parts and assuming that  $\lim_{Z \rightarrow \infty} \varphi = 0$  and  $\lim_{Z \rightarrow \infty} \frac{\partial \varphi}{\partial Z} = 0$ , we get:

$$\begin{aligned} \frac{\partial C}{\partial T} &= -rC + \frac{1}{2} e^{-r(T-t)} \sigma(K, T)^2 K^2 \varphi + r e^{-r(T-t)} \int_K^\infty Z \varphi dZ \\ &= -rC + \frac{1}{2} e^{-r(T-t)} \sigma(K, T)^2 K^2 \varphi \\ &\quad + \underbrace{r e^{-r(T-t)} \int_K^\infty (Z - K) \varphi dZ}_{=C} + \underbrace{rK e^{-r(T-t)} \int_K^\infty \varphi dZ}_{=-\frac{\partial C}{\partial K}} \\ &\stackrel{(2.1.3)}{=} -rC + \frac{1}{2} \sigma(K, T)^2 K^2 + rC - rK \frac{\partial C}{\partial K} \end{aligned}$$

So we finally get the following expression:

$$\frac{\partial C}{\partial T} = \frac{1}{2} \sigma(K, T)^2 K^2 - rK \frac{\partial C}{\partial K} . \quad (2.1.7)$$

Note that in this equation  $\sigma$  is evaluated at  $K$  and  $T$  which is due to the Fokker-Planck equation and the partial integration with lower boundary  $K$ .

By resolving equation (2.1.7) to  $\sigma$  we get the formula of Dupire:

**Theorem 2.2 (Formula of Dupire (1993)).**

Let the asset price process evolve according to the risk-neutral diffusion

$$dS(t) = S(t) [rdt + \sigma(S(t), t) dW^Q(t)]$$

and assume that no dividends are paid.

If today's market call prices  $C(K, T)$  are known for all possible strikes  $K$  and all maturities  $T$ , i.e. there is a continuous set of today's market call prices  $(C(K, T))_{K, T \in \cdot}$ , then we must choose:

$$\sigma(K, T) = \sqrt{\frac{\frac{\partial C}{\partial T} + rK \frac{\partial C}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}}} . \quad (2.1.8)$$

**Proof:**

See the calculations above. ■

**Remark.** Note that this gives  $\sigma(K, T)$  and by relabeling the variables  $\sigma(S(t), t)$ :

$$\sigma(S(t), t) = \sqrt{\frac{\frac{\partial C}{\partial T}|_{T=t} + rS(t) \frac{\partial C}{\partial K}|_{K=S(t)}}{\frac{1}{2} S(t)^2 \frac{\partial^2 C}{\partial K^2}|_{K=S(t)}}} . \quad (2.1.9)$$

**Advantages:**

- Since we still have a 1-D diffusion process which drives the asset price dynamics and therefore “ # stocks = # sources of risk ” and since by assumption  $\sigma > 0$ , the EMM  $Q$  is still unique.
- The uniqueness of  $Q$  implies that the market is still complete and so every contingent claim is attainable.

**Disadvantages:**

- We cannot empirically fulfill the assumptions of theorem 2.2 because no continuous set  $(C(K, T))_{K, T \in \cdot}$  of market prices of call options is available.
- There exists no closed-form solution for European call prices ( $\rightarrow$  we have to use a numerical scheme for the calibration).

**2.2 Calibration: A practical Approach**

First of all we want to explain what we mean when we speak of the calibration of a model. This will be done in the following digression:

**Digression 2: Calibrating a Model**

The calibration of an (equity price) model means estimating its unknown parameters. We need two ingredients:

- Some market prices of options.
- A closed-form solution (or a numerical method) for calculating the corresponding model option prices. These prices depend on some parameters which have to be estimated.

The problem now is:

*Choose the parameters of the model such that the deviation between observed market prices and the theoretical prices is as small as possible.*

**Remark.** As a numerical method one often uses Monte-Carlo simulation for calculating the model prices and for the minimization of the deviation between observed and calculated prices in many cases a least-squared error method is used.

Example: (Calibration of the Black-Scholes model)

The only unknown parameter in the Black-Scholes model is the volatility so this is the only parameter which has to be estimated.

- Let  $(C_M(K_i, T_j))_{i=1, \dots, N; j=1, \dots, M}$  denote the set of available market prices of European calls with strikes  $(K_i)_{i=1, \dots, N}$  and maturities  $(T_j)_{j=1, \dots, M}$  at time  $t = 0$ .
- Let  $C(0, S(0), \sigma, r, K_i, T_j)$  denote the theoretical Black-Scholes price of a European call with strike  $K_i$  and maturity  $T_j$  at time  $t = 0$ .

$\Rightarrow$  Calibration task: Determine the value  $\sigma^*$  solving

$$\min_{\sigma > 0} \sum_{i=1}^N \sum_{j=1}^M (C_M(K_i, T_j) - C(0, S(0), \sigma, r, K_i, T_j))^2 \quad . \quad (2.2.1)$$

(End of Digression 2)

Due to the lack of “all” option prices, the calculation of the derivatives in the Dupire formula might be impossible. One possibility of solving this problem was explained by [Tavella & Klopfer, 2001] who suggested a parametrization of the local volatility surface (i.e. assume a parametric form for the local volatility function  $\sigma(S(t), t)$ ) and calibrated the model by using a numerical solution of the Fokker-Planck equation.

Since there is no closed-form solution for European calls, we have to calibrate the local volatility model by using a numerical method. We know that the price of a European call is given by:

$$C(K, T) = e^{-r(T-t)} \int_K^\infty (Z - K) \varphi(z, t; Z, T) dZ \quad , \quad (2.2.2)$$

where the transition probability density function  $\varphi(z, t; Z, T)$  evolves according to the Fokker-Planck equation (2.1.6) (remember  $z = S(t)$  and  $Z = S(T)$ ).

So if we *solve* equation (2.1.6) *numerically* (which means solving simple initial value problems whose numerical solutions are well-behaved), we can calibrate the local volatility model under the use of equation (2.2.2).

Additionally, since the market option prices are only known for a small number of strikes and maturities, formula (2.1.9) is of limited practical value. The presence of the curvature term in the denominator will cause the shape of the (local) volatility surface to be sensitive to any interpolation device used to compute the partial derivatives in this equation. To avoid this the local volatility function will be *parametrized* in terms of a small number of parameters and a sequence of *optimization* problems will be solved to determine those parameters.

In detail, the procedure is as follows:

**Parametrization:**

To parametrize the local volatility function the following representation is used:

$$\sigma(S(t), t) = \sigma(S(t); \alpha_1(t), \alpha_2(t), \dots, \alpha_5(t)) \quad , \quad (2.2.3)$$

where the  $(\alpha_k)_{k=1, \dots, 5}$  are assumed to be stepwise linear functions of time.

Thereby the solution is not parametrized as a function of time; the time dependence enters implicitly.

In particular we assume following form of the local volatility function:

$$\begin{aligned} \sigma(S(t), t) &= \alpha_1(t) + \frac{\alpha_2(t)}{S(t)^{\alpha_5(t)}} + \alpha_3(t)S(t)^{\alpha_4(t)} \quad , \text{ where} \\ \alpha_k(t) &= \alpha_k(T_{j-1}) + \frac{\alpha_k(T_j) - \alpha_k(T_{j-1})}{T_j - T_{j-1}}(t - T_{j-1}) \quad \text{for } k = 1, \dots, 5. \end{aligned} \quad (2.2.4)$$

Thereby the unknown values  $\alpha_1(T_j), \dots, \alpha_5(T_j)$  as well as  $\alpha_1(T_{j-1}), \dots, \alpha_5(T_{j-1})$  are solutions to the optimization problem explained below.

**Remarks.**

- In the parametrized form, the term  $\alpha_1(t)$  reflects a constant component of the local volatility, the term  $\frac{\alpha_2(t)}{S(t)^{\alpha_5(t)}}$  reflects the increase of volatility as  $S(t)$  decreases and the term  $\alpha_3(t)S(t)^{\alpha_4(t)}$  reflects the more gradual decrease in volatility as  $S(t)$  increases.
- Of course, other functional forms than (2.2.4) like spline functions or Fourier decompositions are possible for the parametrization.

**Optimization:**

To determine the parameters we solve a *sequence* of low-dimensional optimization problems. The advantages of splitting the optimization in a sequence of low-dimensional optimization problems in increasing maturity are that this is much less effort than solving one large-dimensional optimization problem and that there is a great deal of reuse of the computed results.

A simple form of an optimization problem to imply the local volatility function would be a least-squared error problem similar to equation (2.2.1) which minimizes the squared difference between the option prices with the assumed form (2.2.4) of the local volatility function and the observed market prices.

But due to the lack of a continuous set of market prices of options, an interpolation of the partial derivatives of the solution is needed. This would make the solution a discontinuous function which is not smooth.

Since in our case we parametrize the local volatility function as a function of the underlying asset price, we have only to consider the smoothness of the partial derivative of the local volatility function w.r.t. time.

To address this, we add to the simplest optimization problem which would mean a least-squared error problem as explained above, a *smoothing term for the partial derivative of the local volatility function*. To control the influence of this smoothing term on the optimization problem we use a penalty factor  $\lambda$ . The higher we choose  $\lambda$  the smoother looks the computed volatility surface.

To formulate the optimization problem, we need some notation:

- $(C_M(K_i, T_j))_{i=1, \dots, N; j=1, \dots, M}$  again denotes the set of available market prices of European calls with strikes  $(K_i)_{i=1, \dots, N}$  and maturities  $(T_j)_{j=1, \dots, M}$ .
- $\hat{C}(K_i, T_j; \sigma(S(T_j); \alpha_1(T_j), \dots, \alpha_5(T_j)))$  denotes the value of European call with strike  $K_i$  and maturity  $T_j$  computed by numerically solving equation (2.2.2) using the numerical solution of the Fokker-Planck equation (2.1.6), whereby the parametrized form  $\sigma(S(T_j); \alpha_1(T_j), \dots, \alpha_5(T_j))$  of the local volatility function is used to compute this numerical solution.

So we finally have to solve the following sequence of optimization problems:

For  $j = 1, \dots, M$  do:

$$\min_{\alpha_1(T_j), \dots, \alpha_5(T_j)} \left( \lambda \frac{\partial \sigma(S(T_j); \alpha_1(T_j), \dots, \alpha_5(T_j))}{\partial t} + \sum_{i=1}^N \left\{ C_M(K_i, T_j) - \hat{C}(K_i, T_j; \sigma(S(T_j); \alpha_1(T_j), \dots, \alpha_5(T_j))) \right\}^2 \right) \quad (2.2.5)$$

**Remark.** The initial guess for the first optimization problem thereby is determined by trial and error. Each further optimization problem uses the solution of the previous one as a starting guess. Therefore the procedure is quite effective (although the number of optimizations solved is very large), because each optimization starts out with a reasonably good guess.

### **Numerical solution of the FPE:**

During the optimization problem (2.2.5) we have to solve the Fokker-Planck equation corresponding to the parametrized form  $\sigma(S(T_j); \alpha_1(T_j), \dots, \alpha_5(T_j))$  of the local volatility function for maturities  $(T_j)_{j=1, \dots, M}$  numerically. This means solving initial value problems  $\rightarrow$  we need *initial conditions* and *boundary conditions*. Here the initial value problems are solved using a semi-implicit Crank-Nicholson discretization scheme.

#### *Initial conditions:*

In general it holds that if the process  $S$  has the known value  $S(t)$  at the initial time  $t$  the initial condition of the FPE is  $\varphi(S(t), t; S(t'), t) = \delta(S(t') - S(t))$  (see [Wilmott, 2006]). If the process is not known at time  $t$ , but its probability density function  $\varphi(S(t), t; S(t'), t)$  is known, then this is the initial condition.

In the first optimization problem the initial condition will be a delta function since  $S(0)$  is known. For the subsequent optimization problems the initial conditions will be the probability density function at the maturity of the previous problem. So we get altogether following initial conditions:

- First optimization problem:  $T_0 \rightarrow T_1$   
To be computed:  $\varphi(S(0), 0; S(T_1), T_1)$   
Initial condition:  $\varphi(S(0), 0; S(T_1), 0) = \delta(S(T_1) - S(0))$
- Second optimization problem:  $T_1 \rightarrow T_2$   
To be computed:  $\varphi(S(T_1), T_1; S(T_2), T_2)$   
Initial condition:  $\varphi(S(T_1), T_1; S(T_2), T_1)$
- $\vdots$

- M-th optimization problem:  $T_{M-1} \rightarrow T_M$   
 To be computed:  $\varphi(S(T_{M-1}), T_{M-1}; S(T_M), T_M)$   
 Initial condition:  $\varphi(S(T_{M-1}), T_{M-1}; S(T_M), T_{M-1})$

*Boundary conditions:*

The correct formulation of boundary conditions in a finite difference framework as the Crank-Nicholson approach is important because it influences the stability and the speed of the scheme. There are two possibilities that work equally well:

- Pure convection at the boundary:  
 We assume that we can neglect the diffusion terms at the boundaries and solve the following equation:

$$\frac{\partial \varphi}{\partial T} = -\frac{\partial}{\partial S(T)}(rS(T)\varphi) \quad .$$

- Zero curvature at the boundary:

$$\frac{\partial^2 \varphi}{\partial S(T)^2} = 0 \quad .$$

**Summary:**

The procedure described above based on numerically solving the Fokker-Planck equation and on solving a sequence of optimization problems can be applied to any type of derivative. Moreover, the procedure is reasonably fast since once  $\sigma(S(t), t)$  has been computed it can be stored and be used for the pricing of options.

### 2.3 The CEV Model as a local Volatility Model

The constant elasticity of variance (CEV) model assumes that the stock price is driven by the following diffusion process:

$$dS(t) = \mu S(t)dt + \sigma S(t)^\alpha dW(t) \quad ; \quad S(0) = s_0 \geq 0. \quad (2.3.1)$$

**Remarks.**

- From equation (2.3.1) one can easily obtain that the CEV model incorporates a local volatility given by the relation

$$\sigma(S(t), t) = \sigma S(t)^{\alpha-1} \quad . \quad (2.3.2)$$

- The name of the model comes from the following fact:  
 Define the **elasticity of variance** w.r.t. the stock price  $\epsilon_{\sigma(S(t), t)^2, S(t)}$  - which is a measure for the response of variance to price changes - as the relative change in variance divided by the relative price change, then we get:

$$\begin{aligned} \epsilon_{\sigma(S(t), t)^2, S(t)} &= \frac{d\sigma(S(t), t)^2 / \sigma(S(t), t)^2}{dS(t) / S(t)} = \frac{d\sigma(S(t), t)^2}{dS(t)} \cdot \frac{S(t)}{\sigma(S(t), t)^2} \\ &\stackrel{(2.3.2)}{=} \frac{d(\sigma^2 S(t)^{2(\alpha-1)})}{dS(t)} \cdot \frac{S(t)}{\sigma^2 S(t)^{2(\alpha-1)}} = \sigma^2 2(\alpha-1) S(t)^{2\alpha-3} \cdot \frac{S(t)}{\sigma^2 S(t)^{2\alpha-2}} \\ &= 2(\alpha-1) = \text{constant} . \end{aligned}$$

By the deterministic relationship (2.3.2) between stock price and volatility the Black-Scholes model is naturally extended and implied volatility smiles can be fitted:

- If  $\alpha < 1$  then the volatility decreases as the stock price increases which covers the case of a volatility skew in equity markets.
- If  $\alpha > 1$  then the volatility increases as the stock price increases which is observed as a volatility frown in commodity markets.
- If  $\alpha = 0$  then the volatility is constant w.r.t. the stock price and we get the log-normal distribution of the Black-Scholes model.

In the following we restrict ourselves to the case

$$0 < \alpha < 1. \quad (2.3.3)$$

Under condition (2.3.3), the point 0 is an attainable state (see [Delbaen & Shirakawa, 2002]) and as soon as  $S(t)$  reaches zero we keep it equal to zero. Thus the point 0 becomes the absorbing state for the price process  $S(t)$ . The reason why the point 0 is treated in that way is that this is the only way to treat 0 if we want the SDE (2.3.1) to be fulfilled.

A closed-form solution for European calls was first derived by [Cox, 1975] using the probability transition density function for the stock price. [Schroder, 1989] then expressed this formula by the non-central chi-square distribution. We will derive the closed-form solution using the properties of squared Bessel processes following the more recent work of [Delbaen & Shirakawa, 2002].

### 2.3.1 Weak Solutions by Squared Bessel Processes

To derive a closed-form solution for European calls we will first show that some power of a scaled and time-transformed squared Bessel process is a solution to the CEV diffusion (2.3.1). To do so, we need of course the definition of a squared Bessel process:

#### Definition 2.3:

For every  $\delta \geq 0$  and  $x \geq 0$ , the unique strong solution to the equation

$$X(t) = x + \delta t + 2 \int_0^t \sqrt{|X(s)|} dW(s) \quad (2.3.4)$$

is called a  **$\delta$ -dimensional squared Bessel process** started at  $x$ .

We then denote by  $X^{(\delta)}(t) = \{X^{(\delta)}(t)\}_{t \geq 0}$  a squared Bessel process of dimension  $\delta$ .

**Remark.** A squared Bessel process of dimension  $\delta$  obviously follows the stochastic differential equation

$$dX^{(\delta)}(t) = 2\sqrt{|X^{(\delta)}(t)|} dW(t) + \delta dt \quad ; \quad X^{(\delta)}(0) = x. \quad (2.3.5)$$

Now let  $\{X^{(\delta)}(t)\}_{t \geq 0}$  be a squared Bessel process of dimension  $\delta$  and define the first passage time  $\zeta$  of the point 0 for this process by

$$\zeta = \inf \left\{ t > 0; X^{(\delta)}(t) = 0 \right\} \quad .$$

For the parameters  $\nu > 0$  and  $\delta < 2$  we define a deterministic time transformation by

$$\tau_t^{(\delta, \nu)} = \frac{\sigma^2}{2\nu(2-\delta)} \left( 1 - e^{\frac{-2\nu}{2-\delta} \cdot t} \right) \quad .$$

**Lemma 2.4:**

The process  $\{Y^{(\delta,\nu)}(t)\}_{t \geq 0}$  defined by

$$Y^{(\delta,\nu)}(t) = e^{\nu t} \cdot \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta}, \quad (2.3.6)$$

where  $\tau_t^{(\delta,\nu)} \wedge \zeta = \min(\tau_t^{(\delta,\nu)}, \zeta)$ , fullfills

$$dY^{(\delta,\nu)}(t) = \begin{cases} \nu Y^{(\delta,\nu)}(t) dt + \sigma \cdot (Y^{(\delta,\nu)}(t))^{\frac{1-\delta}{2-\delta}} dW^{(\delta,\nu)}(t) & \text{if } \tau_t^{(\delta,\nu)} \leq \zeta, \\ 0 & \text{if } \tau_t^{(\delta,\nu)} > \zeta, \end{cases} \quad (2.3.7)$$

where  $W^{(\delta,\nu)}(t)$  is defined by

$$W^{(\delta,\nu)}(t) = \int_0^{\tau_t^{(\delta,\nu)}} \frac{2-\delta}{\sqrt{\sigma^2 - 2\nu(2-\delta) \cdot u}} dW(u). \quad (2.3.8)$$

**Proof:**

Applying the product rule to equation (2.3.6) yields:

$$dY^{(\delta,\nu)}(t) = \underbrace{e^{\nu t} \cdot d \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta}}_{(*)} + \underbrace{\left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta} \cdot de^{\nu t}}_{(**)}$$

We get for (\*\*):

$$(**) = \nu e^{\nu t} \cdot \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta} dt \stackrel{(2.3.6)}{=} \begin{cases} \nu Y^{(\delta,\nu)}(t) dt & \text{if } \tau_t^{(\delta,\nu)} \leq \zeta. \\ 0 & \text{if } \tau_t^{(\delta,\nu)} > \zeta. \end{cases}$$

Applying Ito's formula yields for (\*):

$$\begin{aligned} (*) &= e^{\nu t} \cdot d \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta} \\ &= e^{\nu t} \underbrace{\frac{\partial \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta}}{\partial X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta)} dX^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta)}_{(\circ)} \\ &\quad + e^{\nu t} \cdot \underbrace{\frac{1}{2} \frac{\partial^2 \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta}}{\partial X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta)^2} < X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) >_{\tau_t^{(\delta,\nu)} \wedge \zeta}}_{(\diamond)} \end{aligned}$$

Calculating  $(\circ)$  and  $(\diamond)$  yields:

$$\begin{aligned} (\circ) &= \left(1 - \frac{1}{2}\delta\right) \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{-\frac{1}{2}\delta} dX^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \\ &\stackrel{(2.3.5)}{=} \left(1 - \frac{1}{2}\delta\right) \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{-\frac{1}{2}\delta} \cdot \left( 2\sqrt{|X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta)|} dW(\tau_t^{(\delta,\nu)} \wedge \zeta) + \delta d\tau_t^{(\delta,\nu)} \wedge \zeta \right) \\ (\diamond) &= -\frac{1}{4}\delta \left(1 - \frac{1}{2}\delta\right) \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{-1-\frac{1}{2}\delta} d < X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) >_{\tau_t^{(\delta,\nu)} \wedge \zeta} \\ &= -\delta \left(1 - \frac{1}{2}\delta\right) \left( X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) \right)^{-\frac{1}{2}\delta} d\tau_t^{(\delta,\nu)} \wedge \zeta, \\ &\text{since } d < X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) >_{\tau_t^{(\delta,\nu)} \wedge \zeta} = 4X^{(\delta)}(\tau_t^{(\delta,\nu)} \wedge \zeta) d\tau_t^{(\delta,\nu)} \wedge \zeta. \end{aligned}$$



Combining the results for (◊) and (◇) then yields for (\*):

$$(*) = e^{\nu t} \cdot (2 - \delta) \left( X^{(\delta)}(\tau_t^{(\delta, \nu)} \wedge \zeta) \right)^{\frac{1}{2} - \frac{1}{2}\delta} dW(\tau_t^{(\delta, \nu)} \wedge \zeta)$$

A side calculation shows that

$$\left( Y^{(\delta, \nu)}(t) \right)^{\frac{1-\delta}{2-\delta}} = e^{\nu t \left( \frac{1-\delta}{1-\delta} - 1 \right)} \cdot e^{\nu t} \cdot \left( X^{(\delta)}(\tau_t^{(\delta, \nu)} \wedge \zeta) \right)^{\frac{1}{2} - \frac{1}{2}\delta},$$

so that we get for (\*):

$$(*) = \left( Y^{(\delta, \nu)}(t) \right)^{\frac{1-\delta}{2-\delta}} \frac{2 - \delta}{e^{\nu t \left( \frac{1-\delta}{1-\delta} - 1 \right)}} dW(\tau_t^{(\delta, \nu)} \wedge \zeta) = \left( Y^{(\delta, \nu)}(t) \right)^{\frac{1-\delta}{2-\delta}} \frac{2 - \delta}{e^{\frac{-\nu}{2-\delta} \cdot t}} dW(\tau_t^{(\delta, \nu)} \wedge \zeta).$$

Another side calculation using the definition of the time transformation  $\tau_t^{(\delta, \nu)}$  yields

$$e^{\frac{-\nu}{2-\delta} \cdot t} = \sqrt{1 - \tau_t^{(\delta, \nu)} \frac{2\nu(2-\delta)}{\sigma^2}},$$

so that (\*) can be written as

$$\begin{aligned} (*) &= \left( Y^{(\delta, \nu)}(t) \right)^{\frac{1-\delta}{2-\delta}} \frac{2 - \delta}{\sqrt{1 - \tau_t^{(\delta, \nu)} \frac{2\nu(2-\delta)}{\sigma^2}}} dW(\tau_t^{(\delta, \nu)} \wedge \zeta) \\ &= \sigma \cdot \left( Y^{(\delta, \nu)}(t) \right)^{\frac{1-\delta}{2-\delta}} \frac{2 - \delta}{\sqrt{\sigma^2 - 2\nu(2-\delta)\tau_t^{(\delta, \nu)}}} dW(\tau_t^{(\delta, \nu)} \wedge \zeta). \end{aligned}$$

Using the definition (2.3.8) of  $W^{(\delta, \nu)}(t)$  we get:

$$(*) = \begin{cases} \sigma \cdot \left( Y^{(\delta, \nu)}(t) \right)^{\frac{1-\delta}{2-\delta}} dW^{(\delta, \nu)}(t) & \text{if } \tau_t^{(\delta, \nu)} \leq \zeta. \\ 0 & \text{if } \tau_t^{(\delta, \nu)} > \zeta. \end{cases}$$

Combining the results for (\*) and (\*\*) gives the desired result (2.3.7). ■

**Lemma 2.5:**

The process  $\{W^{(\delta, \nu)}(t)\}_{t \geq 0}$  defined by (2.3.8) is a Brownian motion.

**Proof:**

By Levy's characterization of Brownian motion we have to show that  $\{W^{(\delta, \nu)}(t)\}_{t \geq 0}$  is a continuous local martingale with  $\langle W^{(\delta, \nu)} \rangle_t = t$ .

We have by definition that

$$dW^{(\delta, \nu)}(t) = \frac{2 - \delta}{\sqrt{\sigma^2 - 2\nu(2-\delta) \cdot \tau_t^{(\delta, \nu)}}} dW(\tau_t^{(\delta, \nu)}).$$

So  $\{W^{(\delta,\nu)}(t)\}_{t \geq 0}$  is a continuous local martingale by the martingale representation theorem.

$$\begin{aligned}
\langle W^{(\delta,\nu)} \rangle_t &= \int_0^{\tau_t^{(\delta,\nu)}} \frac{(2-\delta)^2}{\sigma^2 - 2\nu(2-\delta) \cdot u} du = \left[ \frac{-(2-\delta)^2}{2\nu(2-\delta)} \ln(\sigma^2 - 2\nu(2-\delta) \cdot u) \right]_0^{\tau_t^{(\delta,\nu)}} \\
&= -\frac{(2-\delta)}{2\nu} \left( \ln(\sigma^2 - 2\nu(2-\delta) \cdot \tau_t^{(\delta,\nu)}) - \ln(\sigma^2) \right) \\
&\downarrow \quad \tau_t^{(\delta,\nu)} = \frac{\sigma^2}{2\nu(2-\delta)} \left( 1 - e^{-\frac{2\nu}{2-\delta} \cdot t} \right) \\
&= -\frac{(2-\delta)}{2\nu} \underbrace{\left( \ln(\sigma^2 - \sigma^2 \cdot (1 - e^{-\frac{2\nu}{2-\delta} \cdot t})) - \ln(\sigma^2) \right)}_{=\ln(\sigma^2 \cdot e^{-\frac{2\nu}{2-\delta} \cdot t})} = -\frac{(2-\delta)}{2\nu} \left( \ln(e^{-\frac{2\nu}{2-\delta} \cdot t}) \right) = t
\end{aligned}$$

This completes the proof. ■

**Remark.** If we now choose

$$\delta_\alpha = \frac{1-2\alpha}{1-\alpha} \quad \left( \Rightarrow \frac{1-\delta_\alpha}{2-\delta_\alpha} = \alpha \right) \quad \text{and} \quad \nu = \mu \quad , \quad (2.3.9)$$

we can conclude by equation (2.3.7) that  $Y^{(\delta_\alpha, \mu)}$  fullfills the diffusion process (2.3.1) of the CEV model. Since we restricted ourselves to  $\alpha \in (0, 1)$  we have by  $\delta_\alpha = \frac{1-2\alpha}{1-\alpha}$  that

$$\delta_\alpha \in (-\infty, 1) \quad . \quad (2.3.10)$$

We can summarize the above discussions in the following theorem:

**Theorem 2.6 (Squared Bessel Process as Solution of the CEV Diffusion).**

$$\{S(t)\}_{t \geq 0} \stackrel{\text{law}}{=} \left\{ Y^{(\delta_\alpha, \mu)}(t) \right\}_{t \geq 0} \quad ,$$

where  $\stackrel{\text{law}}{=}$  means equivalence in law under the physical measure  $P$ .

### 2.3.2 Existence of a unique equivalent Martingale Measure

First of all, we shortly repeat how a change of measure from the physical measure  $P$  to a risk neutral measure is done in general.

Therefore let  $\{X(t)\}_{t \geq 0}$  be an  $m$ -dimensional progressively measurable stochastic process adapted to the Brownian filtration with

$$\int_0^t X_i^2(s) ds < \infty \quad P - a.s. \quad \forall t \geq 0 \quad , \quad i = 1, \dots, m.$$

Then define the process

$$Z(t, X) := e^{-\sum_{i=1}^m \int_0^t X_i(s) dW_i(s) - \frac{1}{2} \int_0^t \|X(s)\|^2 ds} \quad . \quad (2.3.11)$$

By Ito's formula we get:

$$Z(t, X) = 1 - \sum_{i=1}^m \int_0^t Z(s, X(s)) X_i(s) dW_i(s) \quad . \quad (2.3.12)$$

This implies that  $Z(t, X)$  is a continuous local martingale with  $Z(0, X) = 1$ . Additionally by definition  $Z(t, X)$  is positive, so it is a supermartingale.

If  $Z(t, X)$  is even a martingale, we have that  $[Z(t, X)] = 1 \forall t \geq 0$ .

If this is the case  $\forall T \geq 0$  we can define a probability measure  $Q_T$  on  $\mathcal{F}_T$  by

$$Q_T(A) := [1_A \cdot Z(T, X)] \quad \forall A \in \mathcal{F}_T, \quad (2.3.13)$$

i.e. we have that the Radon-Nikodym derivative of  $Q_T$  w.r.t.  $P$  is given by  $Z(T, X)$ .

In particular we have by definition that  $Z(T, X)$  is positive which implies that  $P$  and  $Q_T$  are equivalent since they have the same null sets.

Finally we state the well-known theorem of Girsanov which is important to give the stock price dynamics under the measure  $Q$  defined as above:

**Theorem 2.7 (Girsanov).**

Let the process  $Z(t, X)$  defined by (2.3.11) be a martingale and define the process  $\{(W^Q(t), \mathcal{F}_t)\}_{t \geq 0}$  as

$$W_i^Q(t) := W_i(t) + \int_0^t X_i(s) ds, \quad 1 \leq i \leq m, t \geq 0.$$

Then for every fixed  $T \in [0, \infty)$  the process  $\{(W^Q(t), \mathcal{F}_t)\}_{t \geq 0}$  is an  $m$ -dimensional Brownian motion on  $(\Omega, \mathcal{F}_T, Q_T)$  where the probability measure  $Q_T$  is defined by (2.3.13).

**Proof:**

See [Korn & Korn, 2001], pp.108. ■'

For Brownian models (compare section 1.1)

$$\begin{aligned} dB(t) &= B(t)r(t)dt \quad ; \quad B(0) = 1 \\ dS_i(t) &= S_i(t) \left[ \mu_i(t)dt + \sum_{j=1}^d \sigma_{ij}(t)dW_j(t) \right] \quad ; \quad S_i(0) = s_{i0} > 0; \quad i = 1, \dots, n \end{aligned}$$

one chooses  $X(t) := \theta(t)$ , where the process  $\theta(t)$  is implicitly defined as the solution of the following system of linear equations:  $\sigma(t)\theta(t) = \mu(t) - r(t) \cdot 1$ . Since by assumption  $r$  and  $\mu$  are uniformly bounded and  $\sigma$  is positive definite,  $\theta(t)$  fullfills the condition

$$\int_0^T \|\theta(u)\|^2 du < K \quad (K > 0 = \text{constant}) \quad .$$

This condition implies that  $Z(t, \theta)$  is a martingale (see [Korn & Korn, 2001], pp. 111).

Then by the theorem of Girsanov  $W^Q(t)$  with

$$W_i^Q(t) := W_i(t) + \int_0^t \theta_i(s) ds, \quad t \in [0, T] \quad i = 1, \dots, n$$

is  $Q_T$ -Brownian motion w.r.t. the Brownian filtration, where  $Q_T$  is defined by

$$Q_T(A) = [1_A \cdot Z(T, \theta)] \quad , \quad A \in \mathcal{F}_T.$$

As mentioned above  $P$  and  $Q_T$  are equivalent since they have the same null sets. Further it is well-known that the processes

$$U_i(t) := \frac{S_i(t)}{B(t)} \quad , \quad i = 1, \dots, n$$

are martingales w.r.t.  $Q_T$ .

For this reason  $Q_T$  is said to be an **equivalent martingale measure**.

Let us now consider the CEV model case:

$$\begin{aligned} dB(t) &= rB(t)dt \quad ; \quad B(0) = 1 \\ dS(t) &= \mu S(t)dt + \sigma S(t)^\alpha dW(t) \quad ; \quad S(0) = s_0 \geq 0. \end{aligned} \quad (2.3.14)$$

Analogously to the Brownian model above we now want the process  $U(t)$  defined by

$$U(t) := \frac{S(t)}{B(t)} = e^{-rt} S(t) \quad (2.3.15)$$

to be a martingale w.r.t. the new measure  $Q_T$ .

The following lemma motivates how we have to choose the process  $X(t)$  in order to define an equivalent martingale measure for the CEV model.

**Lemma 2.8:**

The process  $U(t)$  defined by (2.3.15) fullfills

$$dU(t) = \sigma e^{-(1-\alpha)rt} U(t)^\alpha d\tilde{W}(t) \quad , \quad (2.3.16)$$

where the process  $\tilde{W}(t)$  is defined as

$$\tilde{W}(t) := W(t) + \int_0^t \theta S(u)^{1-\alpha} du \quad \text{with} \quad \theta := \frac{\mu - r}{\sigma}. \quad (2.3.17)$$

**Proof:**

The product rule yields:

$$\begin{aligned} dU(t) &= e^{-rt} dS(t) + S(t) d(e^{-rt}) = e^{-rt} (\mu S(t) + \sigma S(t)^\alpha dW(t)) - r e^{-rt} S(t) dt \\ &= \sigma e^{-rt} S(t)^\alpha \left( dW(t) + S(t)^{1-\alpha} \frac{\mu - r}{\sigma} dt \right) \end{aligned}$$

By (2.3.17) we have that  $d\tilde{W}(t) = dW(t) + \frac{\mu - r}{\sigma} S(t)^{1-\alpha} dt$ , so that we get:

$$dU(t) = \sigma e^{-rt} S(t)^\alpha d\tilde{W}(t) = \sigma e^{-rt} e^{\alpha rt} e^{-\alpha rt} S(t)^\alpha d\tilde{W}(t) = \sigma e^{-(1-\alpha)rt} U(t) d\tilde{W}(t)$$

■

Equation (2.3.16) and (2.3.17) suggest to choose

$$\begin{aligned} X(t) &:= \theta \cdot S(t)^{1-\alpha} \quad , \text{i.e.} \\ Z(t, \theta \cdot S^{1-\alpha}) &:= e^{-\theta \int_0^t S(u)^{1-\alpha} dW(u) - \frac{1}{2} \theta^2 \int_0^t S(u)^{2(1-\alpha)} du} \end{aligned} \quad (2.3.18)$$

We then define the probability measure  $\tilde{P}_T$  on  $\mathcal{F}_T$  by<sup>3</sup>

$$\tilde{P}_T(A) := [1_A \cdot Z(T, \theta \cdot S^{1-\alpha})] \quad \forall A \in \mathcal{F}_T. \quad (2.3.19)$$

To finally show the existence of the measure  $\tilde{P}_T$  we necessarily need to verify that  $Z(t, \theta \cdot S^{1-\alpha})$  is a martingale.

---

<sup>3</sup>We denote this measure by  $\tilde{P}_T$  and the corresponding Brownian motion by  $\{\tilde{W}(t)\}_{t \geq 0}$  to clearly distinguish them from the measure  $Q_T$  and the Brownian motion  $\{W(t)\}_{t \geq 0}$  used in the Brownian model.

**Remark.** Since the stock price can hit zero, we have to take the minimum of  $T$  and  $\zeta$  so define the probability measure on  $\mathcal{F}_{T \wedge \zeta}$ .

**Proposition 2.9 (Martingale Property of  $Z(t \wedge \zeta, \theta \cdot S^{1-\alpha})$ ).**

For all  $t < \infty$  the process

$$Z(t \wedge \zeta, \theta \cdot S^{1-\alpha}) := e^{-\theta \int_0^{t \wedge \zeta} S(u)^{1-\alpha} dW(u) - \frac{1}{2} \theta^2 \int_0^{t \wedge \zeta} S(u)^{2(1-\alpha)} du}$$

is a martingale.

**Proof:**

The proof of this property would go beyond the scope of this thesis, so we refer to the proof of theorem 2.3 of [Delbaen & Shirakawa, 2002]. ■

**Corollary 2.10:**

There exists a unique  $P$ -equivalent measure  $\tilde{P}$  on  $\mathcal{F}_T$ .

**Proof:**

Existence follows directly since  $Z(t, \theta \cdot S^{1-\alpha})$  is a martingale. Since in our model we have one source of risk (i.e. one Brownian motion) and one stock we get uniqueness. ■

By the theorem of Girsanov we can then deduce that the process  $\left\{(\tilde{W}(t), \mathcal{F}_t)\right\}_{t \geq 0}$  defined as

$$\tilde{W}(t) := W(t) + \int_0^t \theta \cdot S(u)^{1-\alpha} du, \quad t \geq 0$$

is  $\tilde{P}$ -Brownian motion.

From this we get the  $\tilde{P}$ -dynamics of the stock price:

**Theorem 2.11 (Risk-neutral Dynamics under  $\tilde{P}$ ).**

Under the equivalent measure  $\tilde{P}$  defined by (2.3.19) the dynamics of the stock price in the CEV model are given by

$$dS(t) = rS(t)dt + \sigma S(t)^\alpha d\tilde{W}(t) \quad .$$

**Proof:**

Since  $dW(t) = d\tilde{W}(t) - \frac{\mu-r}{\sigma} S(t)^{1-\alpha} dt$  we get:

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sigma S(t)^\alpha dW(t) = \mu S(t)dt + \sigma S(t)^\alpha d\tilde{W}(t) - \sigma S(t)^\alpha \frac{\mu-r}{\sigma} S(t)^{1-\alpha} dt \\ &= rS(t)dt + \sigma S(t)^\alpha d\tilde{W}(t) \quad . \end{aligned}$$

■

From this theorem and lemma 2.4 we can immediately deduce:

**Corollary 2.12:**

$$\{S(t)\}_{t \geq 0} \text{ under } \tilde{P} \stackrel{\text{law}}{=} \left\{Y^{(\delta_\alpha, r)}(t)\right\}_{t \geq 0} \text{ under } P \quad .$$

**Remark.** The equivalent measure  $\tilde{P}$  is indeed a martingale measure. We will show this in the next subsection using a special property of squared Bessel processes.

### 2.3.3 Closed-form Solution for European Calls

Before giving explicit formulas we first need to state some definitions and some important results for squared Bessel processes.

**Definition 2.13 ( $\chi^2$ -distribution):**

Let  $X_1, X_2, \dots$  be i.i.d.  $\mathcal{N}(0, 1)$ -distributed. Then the random variable  $V = \sum_{i=1}^n X_i^2$  is said to be  $\chi^2$ -distributed with  $n$  degrees of freedom (denoted  $\chi_{(n)}^2$ -distribution) and has the density function

$$f(v; n) = \begin{cases} \frac{1}{2^{\frac{n}{2}}} e^{-\frac{v}{2}} v^{\frac{n}{2}-1} \frac{1}{\Gamma(\frac{n}{2})} & \text{for } v > 0, \\ 0 & \text{for } v < 0, \end{cases}$$

where for  $x > 0$   $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$  denotes the  $\Gamma$ -function.

**Definition 2.14 (Noncentral  $\chi^2$ -distribution):**

Let  $X_1, X_2, \dots$  be independent  $\mathcal{N}(\mu_i, 1)$ -distributed. Then the random variable  $V = \sum_{i=1}^n X_i^2$  is said to be noncentral  $\chi^2$ -distributed with  $n$  degrees of freedom and noncentrality parameter  $m = \sum_{i=1}^n \mu_i^2$  (denoted  $\chi_{(n,m)}^2$ -distribution) and has the density function

$$f(v; n, m) = \begin{cases} \frac{1}{2^{\frac{n}{2}}} e^{-\frac{1}{2}(m+v)} v^{\frac{n}{2}-1} \sum_{i=0}^{\infty} \left(\frac{m}{4}\right)^i \frac{v^i}{i! \Gamma(\frac{n}{2} + i)} & \text{for } v > 0, \\ 0 & \text{for } v < 0, \end{cases}$$

where for  $x > 0$   $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$  denotes the  $\Gamma$ -function.

**Lemma 2.15:**

Let  $\{X^{(\delta)}(t)\}_{t \geq 0}$  be a squared bessel process of dimension  $\delta$  and let  $V$  be a random variable with  $V \sim \chi_{(\delta, \frac{x}{t})}^2$  (i.e.  $n = \delta$  and  $m = \frac{x}{t}$ ).

Then for any  $\delta \in [0, \infty)$  we have

$$X(t)^{(\delta)} \stackrel{\text{law}}{=} t \cdot V, \quad x \geq 0, \quad t > 0, \quad \text{where } X(0)^{(\delta)} = x. \quad (2.3.20)$$

**Proof:**

Consider the Laplace transform of  $V \sim \chi_{(\delta, \frac{x}{t})}^2$ :

$$\begin{aligned} [e^{-\lambda V}] &= \int_{v \geq 0}^\infty e^{-\lambda v} \cdot f(v; n, m) dv = \int_{v \geq 0}^\infty e^{-\lambda v} \frac{1}{2^{\frac{n}{2}}} e^{-\frac{1}{2}(m+v)} v^{\frac{n}{2}-1} \sum_{i=0}^{\infty} \left(\frac{m}{4}\right)^i \frac{v^i}{i! \Gamma(\frac{n}{2} + i)} dv \\ &= \int_{v \geq 0}^\infty \frac{1}{2^{\frac{n}{2}}} e^{-\frac{1}{2}(m+(1+2\lambda)v)} v^{\frac{n}{2}-1} \sum_{i=0}^{\infty} \left(\frac{m}{4}\right)^i \frac{v^i}{i! \Gamma(\frac{n}{2} + i)} dv \\ &= \frac{e^{-\frac{\lambda}{1+2\lambda}m}}{(1+2\lambda)^{\frac{b}{2}}} \underbrace{\int_{v \geq 0}^\infty \frac{1}{2^{\frac{n}{2}}} e^{-\frac{1}{2}(\frac{m}{1+2\lambda}+v)} v^{\frac{n}{2}-1} \sum_{i=0}^{\infty} \left(\frac{m}{4(1+2\lambda)}\right)^i \frac{v^i}{i! \Gamma(\frac{n}{2} + i)} dv}_{=1, \text{ since } f(v; n, m) \text{ is a density and } m \rightarrow \frac{m}{1+2\lambda}} \\ &= \frac{e^{-\frac{\lambda}{1+2\lambda}m}}{(1+2\lambda)^{\frac{b}{2}}}. \end{aligned}$$

On the other hand, the Laplace transform of  $X(t)^{(\delta)}$  is given by (see [Revuz & Yor, 1991]):

$$\left[ e^{\lambda X(t)^{(\delta)}} \right] = \frac{e^{-\frac{\lambda t}{1+2\lambda t} \frac{x}{t}}}{(1+2\lambda t)^{\frac{\delta}{2}}} = \left[ e^{-\lambda t V} \right] \quad , \text{ where } V \sim \chi_{(\delta, \frac{x}{t})}^2 .$$

Since the Laplace transform of a random variable  $V$  with  $V \sim \chi_{(\delta, \frac{x}{t})}^2$  is equal to that of  $X(t)^{(\delta)}$  we get (2.3.20). ■

Since we have by equation (2.3.10) that  $\delta_\alpha \in (-\infty, 1)$  we need the duality result for lemma 2.15.

**Lemma 2.16:**

Let  $\{X^{(\delta)}(t)\}_{t \geq 0}$  be a squared bessel process of dimension  $\delta \in (-\infty, 2)$  and let  $\phi$  be a function such that

$$\lim_{x \downarrow 0} \left[ \phi \left( X(t)^{(4-\delta)} \right) \cdot \left( X(t)^{(4-\delta)} \right)^{\frac{\delta}{2}-1} | X(0)^{(4-\delta)} = x \right] < \infty .$$

Then for any  $x > 0$  we have

$$\left[ \phi \left( X(t)^{(\delta)} \right) \cdot 1_{\{\zeta > t\}} | X(0)^{(\delta)} = x \right] = x^{1-\frac{\delta}{2}} \cdot \left[ \phi \left( X(t)^{(4-\delta)} \right) \cdot \left( X(t)^{(4-\delta)} \right)^{\frac{\delta}{2}-1} | X(0)^{(4-\delta)} = x \right] . \quad (2.3.21)$$

**Proof:**

See [Yor, 1992]. ■

Using this lemma, we can now show that  $\tilde{P}$  is a martingale measure.

**Theorem 2.17.**

Define the process  $\{U(t)\}_{t \geq 0}$  by

$$U(t) = \frac{S(t)}{B(t)} = e^{-rt} S(t) \quad .$$

Then we have

$$\tilde{P}[U(t) | U(0) = u] = u$$

and therefore the measure  $\tilde{P}$  is a unique equivalent martingale measure.

**Proof:**

By corollary 2.12 and using the fact that  $Y^{(\delta_\alpha, r)}(t) = e^{rt} \left( X^{(\delta_\alpha)}(\tau_t^{(\delta_\alpha, r)} \wedge \zeta) \right)^{1-\frac{1}{2}\delta_\alpha}$  we get:

$$\begin{aligned} \tilde{P}[U(t) | U(0) = u] &= \tilde{P}[e^{-rt} S(t) | S(0) = u] \\ &= \tilde{P}[e^{-rt} \tilde{Y}^{(\delta_\alpha, r)}(t) | \tilde{Y}^{(\delta_\alpha, r)}(0) = u] \\ &= \tilde{P} \left[ \left( \tilde{X}^{(\delta_\alpha)}(\tau_t^{(\delta_\alpha, r)} \wedge \tilde{\zeta}) \right)^{1-\frac{1}{2}\delta_\alpha} \mid \left( \tilde{X}^{(\delta_\alpha, r)}(0) \right)^{1-\frac{1}{2}\delta_\alpha} = u \right] \\ &= \tilde{P} \left[ \left( \tilde{X}^{(\delta_\alpha)}(\tau_t^{(\delta_\alpha, r)}) \right)^{1-\frac{1}{2}\delta_\alpha} \cdot 1_{\{\tilde{\zeta} \geq t\}} \mid \tilde{X}^{(\delta_\alpha, r)}(0) = u^{\frac{2}{2-\delta_\alpha}} \right] \end{aligned}$$

Then we get by lemma 2.16:

$$\begin{aligned} \tilde{P}[U(t) | U(0) = u] &= \left(u^{\frac{2}{2-\delta_\alpha}}\right)^{1-\frac{1}{2}\delta_\alpha} \tilde{P}\left[\left(\tilde{X}^{(4-\delta_\alpha)}(\tau_t^{(4-\delta_\alpha),r})\right)^{1-\frac{1}{2}\delta_\alpha}\right. \\ &\quad \left.\cdot \left(\tilde{X}^{(4-\delta_\alpha)}(\tau_t^{(\delta_\alpha),r})\right)^{\frac{1}{2}\delta_\alpha-1} \mid \tilde{X}^{(4-\delta_\alpha,r)}(0) = u^{\frac{2}{2-\delta_\alpha}}\right] \\ &= u \end{aligned}$$

■

Now we derive the probability distribution of  $S(T)$  under the physical measure  $P$  by using theorem 2.6, lemma 2.15 and lemma 2.16.

**Theorem 2.18 (Probability Distribution of  $S(T)$ ).**

We can express the probability distribution of the stock price at time  $T$  by

$$P[S(T) \leq x | S(0) = s] = 1 - \sum_{i=1}^{\infty} g(i + \lambda, z) G(i, w) \quad , \quad (2.3.22)$$

where

$$\begin{aligned} \lambda &= \frac{1}{2(1-\alpha)} \\ z &= \frac{s^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, \mu)}} = \frac{2\mu\lambda e^{\frac{\mu T}{\lambda}} s^{\frac{1}{\lambda}}}{\sigma^2 \cdot \left(e^{\frac{\mu T}{\lambda}} - 1\right)} \\ \omega &= \frac{(e^{-\mu T} x)^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, \mu)}} = \frac{2\mu\lambda x^{\frac{1}{\lambda}}}{\sigma^2 \cdot \left(e^{\frac{\mu T}{\lambda}} - 1\right)} \\ g(u, v) &= \frac{v^{u-1}}{\Gamma(u)} e^{-v} \\ G(u, v) &= \int_{\omega \geq v} g(u, \omega) d\omega \quad . \end{aligned} \quad (2.3.23)$$

**Proof:**

By theorem 2.6, equation (2.3.6) and the fact that  $1 - \frac{1}{2}\delta_\alpha = \frac{2-\delta_\alpha}{2}$  we get:

$$\begin{aligned} P[S(T) \geq x | S(0) = s] &= P[Y^{(\delta_\alpha, \mu)}(T) \geq x | Y^{(\delta_\alpha, \mu)}(0) = s] \\ &= P\left[(X^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)} \wedge \zeta))^{\frac{2-\delta_\alpha}{2}} \geq e^{-\mu T} x \mid (X^{(\delta_\alpha)}(0))^{\frac{2-\delta_\alpha}{2}} = s\right] \\ &= P\left[X^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)} \wedge \zeta) \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}} \mid X^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}}\right] \\ &= \left[1\{X^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}) \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}}\} \cdot 1\{\zeta \geq T\} \mid X^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}}\right] \end{aligned}$$

Using lemma 2.16 with  $\phi = 1\{\cdot\}$  we get:

$$\begin{aligned} P[S(T) \geq x | S(0) = s] &= \underbrace{\left(s^{\frac{2}{2-\delta_\alpha}}\right)^{\frac{2-\delta_\alpha}{2}}}_{=s} \left[1\{X^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}) \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}}\} \right. \\ &\quad \left. \cdot (X^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}))^{\frac{\delta_\alpha}{2}-1} \mid X^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}}\right] \end{aligned}$$



Then lemma 2.15 gives us:

$$\begin{aligned}
& \left[ 1\{X^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}) \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}}\} \cdot (X^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}))^{\frac{\delta_\alpha}{2}-1} \mid X^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= \left( \tau_T^{(\delta_\alpha, \mu)} \right)^{\frac{\delta_\alpha}{2}-1} \cdot \left[ V^{\frac{\delta_\alpha}{2}-1} \cdot 1\{\tau_T^{(\delta_\alpha, \mu)} \cdot V \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}}\} \mid X^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right], \\
&\text{where } V \sim \chi^2_{\left(4-\delta_\alpha, \frac{x}{\tau_T^{(\delta_\alpha, \mu)}}\right)} \\
&= \underbrace{\left( \tau_T^{(\delta_\alpha, \mu)} \right)^{\frac{\delta_\alpha}{2}-1} \cdot \left[ V^{\frac{\delta_\alpha}{2}-1} \cdot 1\{V \geq 2\omega\} \right]}_{(*)}, \quad \text{where } V \sim \chi^2_{(4-\delta_\alpha, 2z)}
\end{aligned}$$

Thereby the last equal sign holds because we have:

$$\begin{aligned}
\frac{x}{\tau_T^{(\delta_\alpha, \mu)}} &= \frac{s^{\frac{2}{2-\delta_\alpha}}}{\tau_T^{(\delta_\alpha, \mu)}}, \quad \text{since } x = X^{(4-\delta)}(0) = s^{\frac{2}{2-\delta_\alpha}} \\
&= \frac{s^{2(1-\alpha)}}{\tau_T^{(\delta_\alpha, \mu)}}, \quad \text{since } \delta_\alpha = \frac{1-2\alpha}{1-\alpha} \\
&= 2z \\
1\{\tau_T^{(\delta_\alpha, \mu)} \cdot V \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}}\} &\stackrel{\text{Def. } \delta_\alpha}{=} 1\left\{V \geq \frac{(e^{-\mu T} x)^{2(1-\alpha)}}{\tau_T^{(\delta_\alpha, \mu)}}\right\} = 1\{V \geq 2\omega\}
\end{aligned}$$

By definition of the expectation and definition 2.14, (\*) can be written as:

$$\begin{aligned}
(*) &= \frac{1}{\left(\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \cdot \int_{v \geq 0} 1\{v \geq 2\omega\} \cdot v^{\frac{\delta_\alpha}{2}-1} \cdot f(v; 4-\delta_\alpha, 2z) dv \\
&= \frac{1}{\left(\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \int_{v \geq 0} 1\{v \geq 2\omega\} v^{\frac{\delta_\alpha}{2}-1} v^{1-\frac{\delta_\alpha}{2}} \frac{1}{2^{2-\frac{\delta_\alpha}{2}}} e^{-z-\frac{1}{2}v} \sum_{i=0}^{\infty} \left(\frac{z}{2}\right)^i \frac{v^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} dv \\
&= \frac{1}{2 \cdot \left(2\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \int_{v \geq 0} 1\{v \geq 2\omega\} e^{-z-\frac{1}{2}v} \sum_{i=0}^{\infty} \left(\frac{z}{2}\right)^i \frac{v^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} dv
\end{aligned}$$

Resolving the indicator function and transforming  $v \rightarrow 2v$  yields:

$$\begin{aligned}
(*) &= \frac{1}{2 \left(2\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \int_{v \geq \omega} e^{-z-v} \sum_{i=0}^{\infty} \left(\frac{z}{2}\right)^i \frac{(2v)^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} d(2v) \\
&= \frac{1}{\left(2\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \int_{v \geq \omega} e^{-z-v} \sum_{i=0}^{\infty} z^i \frac{v^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} dv
\end{aligned}$$

Shifting index from  $i$  to  $i-1$  yields:

$$(*) = \frac{1}{\left(2\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \int_{v \geq \omega} e^{-z-v} \sum_{i=1}^{\infty} \underbrace{z^{i-1}}_{=\Gamma(i)} \frac{v^{i-1}}{\Gamma(i+1-\frac{\delta_\alpha}{2})} dv$$

A side calculation shows that  $i + 1 - \frac{\delta_\alpha}{2} = i + \frac{1}{2(1-\alpha)} = i + \lambda$ . Using this fact an rearranging terms we get:

$$\begin{aligned}
 (*) &= \frac{1}{\left(2\tau_T^{(\delta_\alpha, \mu)}\right)^{1-\frac{\delta_\alpha}{2}}} \sum_{i=1}^{\infty} \frac{z^{i-1}}{\Gamma(i+\lambda)} e^{-z} \int_{v \geq \omega} \frac{v^{i-1}}{\Gamma(i)} e^{-v} dv \\
 &= \frac{1}{\left(2\tau_T^{(\delta_\alpha, \mu)} \cdot z\right)^{1-\frac{\delta_\alpha}{2}}} \sum_{i=1}^{\infty} \underbrace{\frac{z^{i+\lambda-1}}{\Gamma(i+\lambda)}}_{g(i+\lambda, z)} e^{-z} \underbrace{\int_{v \geq \omega} \frac{v^{i-1}}{\Gamma(i)} e^{-v} dv}_{\underbrace{g(i, v)}_{G(i, \omega)}}
 \end{aligned}$$

Using the definition of  $z$  we get:

$$\frac{1}{\left(2\tau_T^{(\delta_\alpha, \mu)} \cdot z\right)^{1-\frac{\delta_\alpha}{2}}} = \frac{1}{\left(s^{2(1-\alpha)}\right)^{1-\frac{\delta_\alpha}{2}}} = \frac{1}{s}, \quad \text{since } 1 - \frac{\delta_\alpha}{2} = \frac{1}{2(1-\alpha)}.$$

So we finally get

$$P[S(T) \geq x \mid S(0) = s] = \sum_{i=1}^{\infty} g(i + \lambda, z) G(i, \omega)$$

and this yields (2.3.22). ■

As a special case of theorem 2.18 we can conclude the probability that  $X^{(\delta)}(t)$  hits the point zero for the dimension  $\delta \in (-\infty, 2)$ .

**Corollary 2.19:**

Let  $\{X^{(\delta)}(t)\}_{t \geq 0}$  a squared Bessel process of dimension  $\delta \in (-\infty, 2)$ . Then we have for  $x \geq 0$ :

$$P\left[X^{(\delta)}(u) = 0 \text{ during } 0 \leq u \leq t \mid X^{(\delta)}(0) = x\right] = 1 - \left(\frac{x}{2t}\right)^{1-\frac{\delta}{2}} \sum_{i=1}^{\infty} \frac{\left(\frac{x}{2t}\right)^{i-1}}{\Gamma(i + 1 - \frac{\delta}{2})} e^{-\frac{x}{2t}}. \quad (2.3.24)$$

**Proof:**

Let w.l.o.g.  $x = 0 (\Rightarrow \omega = 0)$ , then

$$G(i, \omega) = \int_{v \geq 0} \frac{v^{i-1}}{\Gamma(i)} e^{-v} dv = \frac{1}{\Gamma(i)} \underbrace{\int_{v \geq 0} v^{i-1} e^{-v} dv}_{\stackrel{\text{Def.}}{=} \Gamma(i)} = \frac{\Gamma(i)}{\Gamma(i)} = 1.$$

Hence we get from theorem 2.18 with  $S(0) = s = x^{1-\frac{\delta_\alpha}{2}}$ :

$$P\left[S(T) = 0 \mid S(0) = x^{1-\frac{\delta_\alpha}{2}}\right] = 1 - \sum_{i=1}^{\infty} g(i + \lambda, z) \quad (+)$$

On the other hand we can conclude from theorem 2.6:

$$\begin{aligned}
P \left[ S(T) = 0 \mid S(0) = x^{1-\frac{\delta_\alpha}{2}} \right] &= P \left[ Y^{(\delta_\alpha, \mu)}(T) = 0 \mid Y^{(\delta_\alpha, \mu)}(0) = x^{1-\frac{\delta_\alpha}{2}} \right] \\
&= P \left[ e^{\mu T} \left( X^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)} \wedge \zeta) \right)^{1-\frac{\delta_\alpha}{2}} = 0 \mid \left( X^{(\delta_\alpha)}(0) \right)^{1-\frac{\delta_\alpha}{2}} = x^{1-\frac{\delta_\alpha}{2}} \right] \\
&= P \left[ X^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)} \wedge \zeta) = 0 \mid X^{(\delta_\alpha)}(0) = x \right] \\
&= P \left[ \zeta \leq \tau_T^{(\delta_\alpha, \mu)} \mid X^{(\delta_\alpha)}(0) = x \right]
\end{aligned}
\tag{++}$$

Substituting  $\tau_T^{(\delta_\alpha, \mu)} = t$  and  $\delta_\alpha = \delta$  yields on the one hand:

$$(++) = P \left[ \zeta \leq t \mid X^{(\delta)}(0) = x \right] = P \left[ X^{(\delta)}(u) = 0 \text{ during } 0 \leq u \leq t \mid X^{(\delta)}(0) = x \right]$$

On the other hand from  $\tau_T^{(\delta_\alpha, \mu)} = t$  we get by the definition of  $z$ :

$$z = \frac{s^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, \mu)}} = \frac{s^{2(1-\alpha)}}{2t} = \frac{x}{2t}, \quad \text{since } s = x^{1-\frac{\delta_\alpha}{2}} = x^{\frac{1}{2(1-\alpha)}}$$

From this fact and the definition of  $g(i + \lambda)$  we get:

$$\begin{aligned}
(+) &= 1 - \sum_{i=1}^{\infty} \frac{z^{i+\lambda-1}}{\Gamma(i+\lambda)} e^{-z} = 1 - \sum_{i=1}^{\infty} \frac{\left(\frac{x}{2t}\right)^{i+\lambda-1}}{\Gamma(i+\lambda)} e^{-\frac{x}{2t}} = 1 - \left(\frac{x}{2t}\right)^\lambda \sum_{i=1}^{\infty} \frac{\left(\frac{x}{2t}\right)^{i-1}}{\Gamma(i+\lambda)} e^{-\frac{x}{2t}} \\
&= 1 - \left(\frac{x}{2t}\right)^{1-\frac{\delta_\alpha}{2}} \sum_{i=1}^{\infty} \frac{\left(\frac{x}{2t}\right)^{i-1}}{\Gamma(i+1-\frac{\delta_\alpha}{2})} e^{-\frac{x}{2t}}, \quad \text{since } \lambda = \frac{1}{2(1-\alpha)} = 1 - \frac{\delta_\alpha}{2} \\
&= 1 - \left(\frac{x}{2t}\right)^{1-\frac{\delta}{2}} \sum_{i=1}^{\infty} \frac{\left(\frac{x}{2t}\right)^{i-1}}{\Gamma(i+1-\frac{\delta}{2})} e^{-\frac{x}{2t}} \quad \text{by choosing } \delta_\alpha = \delta.
\end{aligned}$$

Combining (+) and (++) we get the desired result. ■

**Remark.** As a consequence of corollary 2.19 we get that the possibility for the price to attain 0 is strictly positive for  $0 < \alpha < 1$  ! Therefore we have to keep  $S$  equal to zero when it once became zero. Otherwise we would allow arbitrage opportunities because we could by the stock when  $S = 0$  and sell as soon as  $S > 0$  after beeing reflected at zero.

Putting all the results together, we can finally derive a closed-form solution for an option whose payoff depends on the stock price at the maturity  $S(T)$ , i.e. the option has a payoff of the form  $C(S(T))$  for some function  $C(\cdot)$ . We use the fact that the unique arbitrage-free price of an option is given by its discounted expected value under the equivalent martingale measure  $\tilde{P}$ .

**Theorem 2.20 (Closed-form Solution for Payoff  $C(S(T))$ ).**

Let the initial stock price be  $S(0) = s$  and denote by  $C(0, s)$  the arbitrage free price for an option at time 0 with payoff  $C(S(T))$ . Then we have:

$$\begin{aligned}
C(0, s) &= s \cdot \tilde{P} \left[ e^{-rT} \cdot C \left( e^{rT} \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \right) \cdot \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} \right. \\
&\quad \left. | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{2(1-\alpha)} \right] + e^{-rT} C(0) \left( 1 - s \left( \frac{1}{2\tau_T^{(\delta_\alpha, r)}} \right)^{1-\frac{\delta_\alpha}{2}} \sum_{i=1}^{\infty} \frac{\left( \frac{s^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, r)}} \right)^{i-1}}{\Gamma(i+1-\frac{\delta_\alpha}{2})} e^{-\frac{s^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, r)}}} \right),
\end{aligned} \tag{2.3.25}$$

where  $\tilde{P}[\cdot]$  means the expectation under the unique equivalent martingale measure  $\tilde{P}$ .

**Proof:**

It is well-known that in a complete market  $C(0, s)$  is given by its discounted expected value under the unique equivalent martingale measure.

So we get:

$$\begin{aligned}
C(0, s) &= \tilde{P} [e^{-rT} \cdot C(S(T)) | S(0) = s] = \tilde{P} [e^{-rT} \cdot C(\tilde{Y}^{(\delta_\alpha, r)}(T)) | \tilde{Y}^{(\delta_\alpha, r)}(0) = s] \\
&= \tilde{P} \left[ e^{-rT} \cdot C \left( e^{rT} \left( \tilde{X}^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)} \wedge \tilde{\zeta}) \right)^{1-\frac{\delta_\alpha}{2}} \right) | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= \tilde{P} \left[ e^{-rT} \cdot C \left( e^{rT} \left( \tilde{X}^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \right) \cdot 1_{\{\tilde{\zeta} > \tau_T^{(\delta_\alpha, r)}\}} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&\quad + \tilde{P} \left[ e^{-rT} C \left( e^{rT} \left( \tilde{X}^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \right) \cdot 1_{\{\tilde{\zeta} \leq \tau_T^{(\delta_\alpha, r)}\}} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= \tilde{P} \left[ e^{-rT} \cdot C \left( e^{rT} \left( \tilde{X}^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \right) \cdot 1_{\{\tilde{\zeta} > \tau_T^{(\delta_\alpha, r)}\}} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&\quad + e^{-rT} \cdot C(0) \cdot \tilde{P} \left[ \tilde{\zeta} \leq \tau_T^{(\delta_\alpha, r)} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right]
\end{aligned}$$

From lemma 2.16 we get:

$$\begin{aligned}
&\tilde{P} \left[ e^{-rT} \cdot C \left( e^{rT} \left( \tilde{X}^{(\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \right) \cdot 1_{\{\tilde{\zeta} > \tau_T^{(\delta_\alpha, r)}\}} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= s \cdot \tilde{P} \left[ e^{-rT} C \left( e^{rT} \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \right) \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right]
\end{aligned}$$

From corollary 2.19 we get:

$$\begin{aligned}
&\tilde{P} \left[ \tilde{\zeta} \leq \tau_T^{(\delta_\alpha, r)} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] = P \left[ \tilde{X}^{(\delta_\alpha)}(u) = 0 \text{ during } 0 \leq u \leq \tau_T^{(\delta_\alpha, r)} | \tilde{X}^{(\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= 1 - s \left( \frac{1}{2\tau_T^{(\delta_\alpha, r)}} \right)^{1-\frac{\delta_\alpha}{2}} \sum_{i=1}^{\infty} \frac{\left( \frac{s^{\frac{2}{2-\delta_\alpha}}}{2\tau_T^{(\delta_\alpha, r)}} \right)^{i-1}}{\Gamma(i+1-\frac{\delta_\alpha}{2})} e^{-\frac{s^{\frac{2}{2-\delta_\alpha}}}{2\tau_T^{(\delta_\alpha, r)}}}
\end{aligned}$$

Substituting this results in the equation for  $C(0, s)$  and using the fact that

$$\frac{2}{2-\delta_\alpha} \stackrel{(2.3.9)}{=} \frac{2}{1-\frac{1-2\alpha}{1-\alpha}} = \frac{2}{1-\alpha} = 2(1-\alpha) \text{ we get (2.3.25). } \blacksquare$$

As a consequence of theorem 2.20 we can deduce a closed-form solution for European Calls with payoff  $C(S(T)) = \max\{S(T) - K; 0\}$ .

**Theorem 2.21 (Closed-form Solution for European Calls).**

Let the initial stock price be  $S(0) = s$  and denote with  $C(0, s) = C(0, s; \alpha, \sigma, r, K)$  the arbitrage-free price of a European call at time 0 with strike price  $K$ . Then we have:

$$C(0, s; \alpha, \sigma, r, K) = s \sum_{i=1}^{\infty} g(i, z') G(i + \lambda, \omega') - e^{-rT} K \sum_{i=1}^{\infty} g(i + \lambda, z') G(i, \omega') \quad , \quad (2.3.26)$$

where

$$\begin{aligned} \lambda &= \frac{1}{2(1-\alpha)} \\ z' &= \frac{s^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, r)}} = \frac{2\mu\lambda e^{\frac{rT}{\lambda}} s^{\frac{1}{\lambda}}}{\sigma^2 \cdot \left(e^{\frac{rT}{\lambda}} - 1\right)} \\ \omega' &= \frac{(e^{-rT} K)^{2(1-\alpha)}}{2\tau_T^{(\delta_\alpha, r)}} = \frac{2r\lambda K^{\frac{1}{\lambda}}}{\sigma^2 \cdot \left(e^{\frac{rT}{\lambda}} - 1\right)} \\ g(u, v) &= \frac{v^{u-1}}{\Gamma(u)} e^{-v} \\ G(u, v) &= \int_{\omega \geq v} g(u, \omega) d\omega \quad . \end{aligned} \quad (2.3.27)$$

**Proof:**

Since for a European call option we have  $C(S(T)) = \max\{S(T) - K; 0\}$  and therefore  $C(S(T) = 0) = 0$  we get by theorem 2.20:

$$\begin{aligned} C(0, s) &= s \cdot \tilde{P} \left[ e^{-rT} \left( e^{rT} \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} - K \right) \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{1-\delta_\alpha}} \right] \\ &= s \cdot \tilde{P} \left[ \left( 1 - e^{-rT} K \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} \right) 1\{e^{rT} \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{1-\frac{\delta_\alpha}{2}} \geq K\} \right. \\ &\quad \left. | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{1-\delta_\alpha}} \right] \\ &= s \cdot \tilde{P} \left[ \left( 1 - e^{-rT} K \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} \right) 1\{\tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \geq (e^{-rT} K)^{\frac{2}{2-\delta_\alpha}}\} \right. \\ &\quad \left. | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{1-\delta_\alpha}} \right] \quad (*) \\ &= s \tilde{P} \left[ \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \geq (e^{-rT} K)^{\frac{2}{2-\delta_\alpha}} | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\ &\quad - s e^{-rT} K \cdot \tilde{P} \left[ \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} 1\{\tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \geq (e^{-rT} K)^{\frac{2}{2-\delta_\alpha}}\} | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \end{aligned}$$

From lemma 2.15 we get:

$$\begin{aligned}
& \tilde{P} \left[ \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \geq (e^{-rT} K)^{\frac{2}{2-\delta_\alpha}} \mid \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= \tilde{P} \left[ \tau_T^{(\delta_\alpha, r)} V \geq (e^{-rT} K)^{\frac{2}{2-\delta_\alpha}} \mid \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right], \text{ where } V \sim \chi_{(4-\delta_\alpha, \frac{x}{\tau_T^{(\delta_\alpha, r)}})}^2 \\
&= \tilde{P} \left[ V \geq 2\omega' \right], \text{ where } V \sim \chi_{(4-\delta_\alpha, 2z')}^2 \quad (+)
\end{aligned}$$

Thereby the last equal sign holds because we have:

$$\begin{aligned}
\frac{(e^{-rT} K)^{\frac{2}{2-\delta_\alpha}}}{\tau_T^{(\delta_\alpha, r)}} &= \frac{(e^{-rT} K)^{2(1-\alpha)}}{\tau_T^{(\delta_\alpha, r)}} =: 2\omega' \\
\frac{x}{\tau_T^{(\delta_\alpha, r)}} &= \frac{s^{\frac{2}{2-\delta_\alpha}}}{\tau_T^{(\delta_\alpha, r)}} = \frac{s^{2(1-\alpha)}}{\tau_T^{(\delta_\alpha, r)}} =: 2z', \text{ since } x = \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}}
\end{aligned}$$

We finally calculate (+):

$$\begin{aligned}
(+) &= \int_{v \geq 0} 1\{v \geq 2\omega'\} f(v; 4 - \delta_\alpha, 2z') dv \\
&= \int_{v \geq 0} 1\{v \geq 2\omega'\} \frac{1}{2^{2-\frac{\delta_\alpha}{2}}} e^{-\frac{1}{2}(2z'+v)} v^{1-\frac{\delta_\alpha}{2}} \sum_{i=0}^{\infty} \left(\frac{z'}{2}\right)^i \frac{v^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} dv \\
&= \frac{1}{2 \left(2^{1-\frac{\delta_\alpha}{2}}\right)} \int_{v \geq \omega'} e^{-z'-v} 2^{1-\frac{\delta_\alpha}{2}} v^{1-\frac{\delta_\alpha}{2}} \sum_{i=0}^{\infty} \left(\frac{z'}{2}\right)^i \frac{(2v)^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} d(2v) \\
&= \int_{v \geq \omega'} e^{-z'-v} v^{1-\frac{\delta_\alpha}{2}} \sum_{i=0}^{\infty} (z')^i \frac{v^i}{i! \Gamma(i+2-\frac{\delta_\alpha}{2})} dv \\
&= \int_{v \geq \omega'} e^{-z'-v} v^{1-\frac{\delta_\alpha}{2}} \sum_{i=1}^{\infty} (z')^{i-1} \underbrace{\frac{v^{i-1}}{(i-1)!}}_{=\Gamma(i)} \Gamma(i+1-\frac{\delta_\alpha}{2}) dv
\end{aligned}$$

Rearranging terms and using  $1 - \frac{\delta_\alpha}{2} = \lambda$  yields:

$$(+) = \sum_{i=1}^{\infty} \underbrace{\frac{(z')^{i-1}}{\Gamma(i)} e^{-z'}}_{g(i, z')} \underbrace{\int_{v \geq \omega'} \frac{v^{i+\lambda-1} e^{-v}}{\Gamma(i+\lambda)} dv}_{g(i+\lambda, v)} = \sum_{i=1}^{\infty} g(i, z') G(i+\lambda, \omega')$$

From the proof of theorem 2.18 we know that

$$\begin{aligned}
& \left[ (X^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}))^{\frac{\delta_\alpha}{2}-1} \cdot 1\{X^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, \mu)}) \geq (e^{-\mu T} x)^{\frac{2}{2-\delta_\alpha}}\} \mid X^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\
&= \frac{1}{s} \sum_{i=1}^{\infty} g(i+\lambda, z) G(i, \omega) \quad .
\end{aligned}$$

So we can deduce:

$$\begin{aligned} & \tilde{P} \left[ \left( \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \right)^{\frac{\delta_\alpha}{2}-1} 1_{\{ \tilde{X}^{(4-\delta_\alpha)}(\tau_T^{(\delta_\alpha, r)}) \geq (e^{-rT} K)^{\frac{2}{2-\delta_\alpha}} \}} | \tilde{X}^{(4-\delta_\alpha)}(0) = s^{\frac{2}{2-\delta_\alpha}} \right] \\ &= \frac{1}{s} \sum_{i=1}^{\infty} g(i + \lambda, z') G(i, \omega') \quad . \quad (++) \end{aligned}$$

Inserting (+) and (++) in (\*) gives (2.3.26). ■

**Remark.** The CEV model will be important later when examining the Bergomi model!

### 3 Heston and its Shortfalls

In this chapter we examine the properties of the Heston model which is a member of the class of stochastic volatility models. These are called **third generation models**. First, we motivate why volatility should be modeled as a stochastic process. The next step is to describe the model structure. Then we provide a (semi)-closed solution for European calls in the Heston model which is done by deriving a pricing equation analogously to the Black-Scholes PDE and solving this equation. Finally we show in which cases the Heston model reaches his limits. To do so, we investigate the example of a Napoleon option, the dynamics of the Heston parameters, the dynamics of implied volatilities and especially the term structure of the volatility of volatility, the pricing of forward-starting options and the local dynamics in the Heston model. This motivates why a new class of models is needed.

#### 3.1 Motivation

In chapter 1, we already examined the stylized facts **volatility clustering** and **leverage effect**. Volatility clustering in particular implies that volatility is autocorrelated and both imply that price changes and volatility changes are (negatively) **correlated**.

Furthermore, volatility is **mean-reverting**, i.e. it tends to return to its **mean-reversion level**. This property can be explained by a simple economic argument:

Consider the distribution of the volatility of a traded stock over a very long time period (about 100 years). If volatility were not mean-reverting - which would in consequence mean that the distribution is not stable - the probability of the volatility of the stock being in the range of 1% – 100% would be rather low. This is clearly a contradiction since we believe that it is overwhelmingly likely that the volatility indeed is in this range.

We get another argument for the mean-reversion property of volatility when we look at historical stock data. A simple way to estimate the volatility is to use the standard deviation of the (log) returns as a measure of it. This is called historic volatility.

**Definition 3.1 (N-day historic Volatility):**

Let  $\{r(n)\}_{n=1,\dots,M}$  denote the time series of the log return of an asset or an index. Then the N-day historic volatility of an asset or an index is for  $M \geq N$  defined as the annualized standard deviation

$$\sigma_{\text{hist}}(N) = \sqrt{\frac{252}{N} \sum_{n=1}^N (r(n) - \bar{r}_N)^2} \quad , \quad (3.1.1)$$

where  $\bar{r}_N = \frac{1}{N} \sum_{n=1}^N r(n)$  is the mean return.

**Remark.** As one works with annualized quantities and usually daily stock price data are used, we needed the factor 252 in (3.1.1) supposing that there are approximately 252 business days in a year.

To be able to say something about the historic volatility we have to observe it over a longer time period. To do so, we define as next what mean by the notion rolling historic volatility.

**Definition 3.2 (Rolling N-day historic Volatility):**

By the N-day historic volatility of an asset or an index we understand the time series



$$\{\sigma_{\text{hist}}(N, l)\}_{l=0, \dots, M-N} = \left\{ \sqrt{\frac{252}{N} \sum_{n=l+1}^{N+l} (r(n) - \bar{r}_N(l))^2} \right\}_{l=0, \dots, M-N}, \quad (3.1.2)$$

where  $\bar{r}_N(l) = \frac{1}{N} \sum_{n=l+1}^{N+l} r(n)$ .

**Remark.** This means that the  $N$ -day interval of the historic volatility slides within the total observation period  $n = 1, \dots, M$ .

In figure 9 we plot the rolling one-year historic volatility  $\{\sigma_{\text{hist}}(252, l)\}_{l=0, \dots, M-252}$  of the S&P 500 and the N225 Index from 01/04/1985 to 04/03/2007. One can observe for both indices that after hitting a maximum (minimum) the historic volatility tends to fall (rise). This gives empirical evidence for the fact that volatility is indeed mean-reverting.

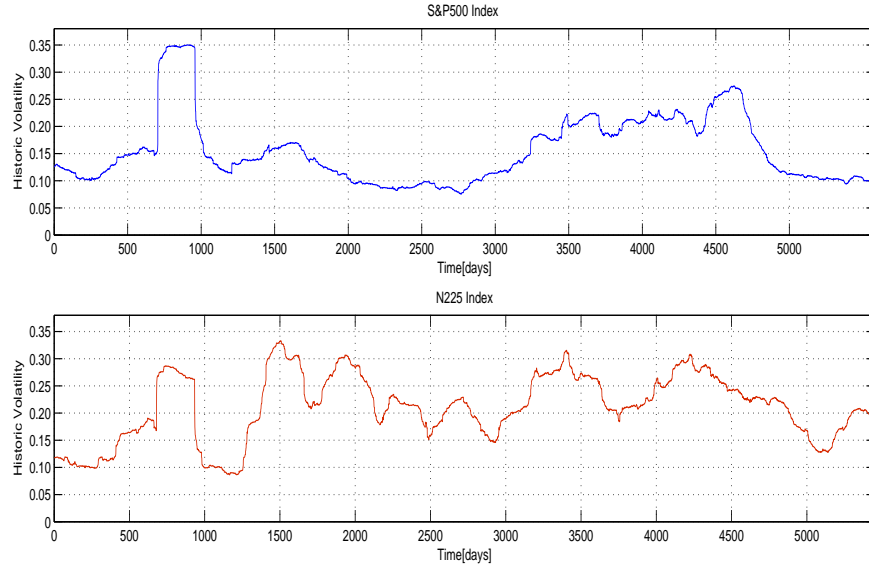


Figure 9: Rolling one-year historic Volatility of S&P 500 Index and N225 Index from 01/04/1985 to 04/13/2007

Putting these facts together motivates to model volatility as a mean-reverting random variable which is correlated with the asset price. This is the basic idea of stochastic volatility models.

### 3.2 Model Structure

The Heston model assumes that the stock price evolves according to the SDE

$$dS(t) = \mu(t)S(t)dt + \sqrt{\nu(t)}S(t)dW_1(t) \quad , \quad (3.2.1)$$

where the volatility itself follows an Ornstein-Uhlenbeck process, i.e.

$$d\sqrt{\nu(t)} = -\beta\sqrt{\nu(t)}dt + \delta dW_2(t) \quad (3.2.2)$$

with

$$\text{Corr}(dW_1, dW_2) = \langle dW_1, dW_2 \rangle = \rho dt. \quad (3.2.3)$$

Applying Ito's formula we can easily deduce according to which process the variance  $\nu(t)$  evolves if the volatility  $\sigma(t) := \sqrt{\nu(t)}$  follows the process (3.2.2).

**Lemma 3.3:**

If the volatility  $\sigma(t) := \sqrt{\nu(t)}$  follows the stochastic differential equation (3.2.2) we have that the variance  $\nu(t)$  evolves according to

$$d\nu(t) = [\delta^2 - 2\beta\nu(t)] dt + 2\delta\sqrt{\nu(t)}dW_2(t) \quad . \quad (3.2.4)$$

**Proof:**

We have that  $\sigma(t) = \sqrt{\nu(t)}$  and therefore  $\sigma(t)^2 = \nu(t)$ . Applying Ito's formula with  $f(\sigma(t)) = \sigma(t)^2 = \nu(t)$  yields:

$$d\nu(t) = d\sigma(t)^2 = \frac{\partial\sigma(t)^2}{\partial\sigma(t)}d\sigma(t) + \frac{1}{2}\frac{\partial^2\sigma(t)^2}{\partial\sigma(t)^2}\underbrace{d\langle\sigma(t)\rangle_t}_{=\delta^2 dt}$$

For the partial derivatives we have:

$$\frac{\partial\sigma(t)^2}{\partial\sigma(t)} = 2\sigma(t) = 2\sqrt{\nu(t)} \quad \text{and} \quad \frac{1}{2}\frac{\partial^2\sigma(t)^2}{\partial\sigma(t)^2} = \frac{1}{2}\frac{\partial}{\partial\sigma(t)}\left(\frac{\partial\sigma(t)^2}{\partial\sigma(t)}\right) = \frac{1}{2}\frac{\partial}{\partial\sigma(t)}(2\sigma(t)) = 1$$

Using this and plugging in the dynamics of the volatility we get:

$$d\nu(t) = 2\sqrt{\nu(t)}\left[-\beta\sqrt{\nu(t)}dt + \delta dW_2(t)\right] + \delta^2 dt = [\delta^2 - 2\beta\nu(t)] dt + 2\delta\sqrt{\nu(t)}dW_2(t) \quad .$$

This completes the proof. ■

If we now choose  $\delta^2 = \kappa\theta$ ,  $2\beta = \kappa$  and  $2\delta = \sigma$  we can represent (3.2.4) as a square-root process:

$$d\nu(t) = \kappa[\theta - \nu(t)] dt + \sigma\sqrt{\nu(t)}dW_2(t) \quad . \quad (3.2.5)$$

Altogether we can summarize the approach by Heston as follows:

The stock price and the variance evolve according to the stochastic differential equations

$$\begin{aligned} dS(t) &= \mu(t)S(t)dt + \sqrt{\nu(t)}S(t)dW_1(t) \\ d\nu(t) &= \kappa[\theta - \nu(t)] dt + \sigma\sqrt{\nu(t)}dW_2(t) \\ \langle dW_1, dW_2 \rangle &= \rho dt. \end{aligned} \quad (3.2.6)$$

Thereby the parameters are as follows:

- $\{W_1(t)\}_{t \geq 0}$  and  $\{W_2(t)\}_{t \geq 0}$  are two correlated Brownian motions with correlation  $\rho$ .
- $\mu(t)$  : instantaneous drift
- $\nu(t)$  : variance
- $\sqrt{\nu(t)}$  : volatility

- $\theta > 0$  : mean reversion level
- $\kappa > 0$  : mean reversion speed
- $\sigma > 0$  : volatility of volatility

The square-root process (3.2.5) is well-known from the Cox-Ingersoll-Ross model for the short rate. It is also well-known from interest-rate theory (see for example [Brigo & Mercurio, 2006]) that a square-root process has the property that it is non-negative, i.e.

$$P(\nu(t) \geq 0) = 1 \quad \forall t \geq 0 \quad ,$$

and if additionally we have that

$$2\kappa\theta \geq \sigma^2 \quad ,$$

the process is even strictly positive, i.e.

$$P(\nu(t) > 0) = 1 \quad \forall t \geq 0 \quad ,$$

which is a necessary requirement for the volatility.

Additionally  $\nu(t)$  covers the mean-reversion property explained in the motivation:

- If in (3.2.5)  $\nu(t) > \theta \Rightarrow \kappa(\theta - \nu(t)) < 0$ .
- If in (3.2.5)  $\nu(t) < \theta \Rightarrow \kappa(\theta - \nu(t)) > 0$ .

This indicates that  $\nu(t)$  oscillates around the mean-reversion level  $\theta$ .

The correlation of  $\{W_1(t)\}_{t \geq 0}$  and  $\{W_2(t)\}_{t \geq 0}$  covers the correlation of the stock price changes and the changes in volatility. For simulation one chooses then  $\rho \cong -0.5$  to  $-0.8$  in order to produce the negative correlation which is given by the leverage effect. So by the Heston model we can model all properties that we mentioned in the motivation by the set of equations (3.2.6).

**Remark.** As the volatility can contain random sources which are different from the random source of the stock price and the volatility is no traded asset we have that “# sources of risk  $>$  # assets”. Therefore the market is incomplete and there exist infinitely many equivalent martingale measures.

### 3.3 Derivation of the Pricing Equation

In order to obtain a closed-form solution for European calls we first derive the pricing equation for the Heston model which we will solve using appropriate boundary conditions. Remember that this is exactly the same strategy as in the Black-Scholes framework, where we derived the Black-Scholes PDE and then obtained the Black-Scholes formula by solving this equation. We follow [Gatheral, 2006].

At first, we derive the pricing equation starting with a more general equation for the variance than (3.2.5). We suppose that the stock price and its variance satisfy the following SDEs:

$$\begin{aligned} dS(t) &= \mu(t)S(t)dt + \sqrt{\nu(t)}S(t)dW_1(t) \\ d\nu(t) &= \alpha(S, \nu, t)dt + \sigma\beta(S, \nu, t)\sqrt{\nu(t)}dW_2(t) \\ < dW_1, dW_2 > &= \rho dt. \end{aligned} \tag{3.3.1}$$

We then get immediately the results for the Heston case if we substitute  $\alpha(S, \nu, t) = \kappa(\theta - \nu(t))$  and  $\beta(S, \nu, t) = 1$ .

Analogously to the derivation of the Black-scholes PDE we construct a riskless portfolio. In the Black-Scholes case we only had the stock price as a source of randomness which can be hedged with stock (remember the delta hedging strategy in theorem 1.2). Now the volatility itself is a random source too, which we additionally need to hedge to construct a riskless portfolio. Therefore assume that there exist price processes

- $V = V(S, \nu, t) \in C^{1,2}$  of the option  $C$  being priced such that the time- $t$  price of the option is given by  $C(t) = V(S(t), \nu(t), t)$  and
- $V_1 = V_1(S, \nu, t)$  of another option  $C_1$  different from  $C$  such that the time- $t$  price of the option is given by  $C_1(t) = V_1(S(t), \nu(t), t)$ .

Consider then the self-financing trading strategy

$$\varphi(t) = (\varphi_V(t), -\varphi_B(t), -\varphi_S(t), -\varphi_{V_1}(t)) = (1, -\varphi_B(t), -\varphi_S(t), -\varphi_{V_1}(t)) \quad , \text{ i.e.}$$

- buy one unit of the option being priced  $C$ ,
- sell  $\varphi_B$  units on the money market account,
- sell  $\varphi_S$  units of the stock  $S$ ,
- sell  $\varphi_{V_1}$  units of the asset  $C_1$ .

Using this trading strategy we get for the wealth of the investor (value of the portfolio):

$$\Pi^\varphi(t) = V(S(t), \nu(t), t) - \varphi_B(t) B(t) - \varphi_S(t) S(t) - \varphi_{V_1}(t) V_1(S(t), \nu(t), t). \quad (3.3.2)$$

In the following we skip the dependences and write

$$\Pi^\varphi = V - \varphi_B B - \varphi_S S - \varphi_{V_1} V_1$$

as short-hand notation.

Thereby we assume for the moment that the interest rate is constant, so the money market account behaves as follows:

$$dB = rBdt.$$

Since the trading strategy  $\varphi$  was assumed to be self-financing and by Ito's formula we get for the dynamics of the portfolio:

$$\begin{aligned} d\Pi^\varphi &= dV - d(\varphi_B B) - d(\varphi_S S) - d(\varphi_{V_1} V_1) = dV - \varphi_B dB - \varphi_S dS - \varphi_{V_1} dV_1 \\ &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial \nu} d\nu - \varphi_B r B dt - \varphi_S dS \\ &\quad + \frac{1}{2} \left[ \frac{\partial^2 V}{\partial S^2} d\langle S \rangle_t + \frac{\partial^2 V}{\partial \nu \partial S} d\langle S, \nu \rangle_t + \frac{\partial^2 V}{\partial S \partial \nu} d\langle \nu, S \rangle_t + \frac{\partial^2 V}{\partial \nu^2} d\langle \nu \rangle_t \right] \\ &\quad - \varphi_{V_1} \left\{ \frac{\partial V_1}{\partial t} dt + \frac{\partial V_1}{\partial S} dS + \frac{\partial V_1}{\partial \nu} d\nu \right. \\ &\quad \left. + \frac{1}{2} \left[ \frac{\partial^2 V_1}{\partial S^2} d\langle S \rangle_t + \frac{\partial^2 V_1}{\partial \nu \partial S} d\langle S, \nu \rangle_t + \frac{\partial^2 V_1}{\partial S \partial \nu} d\langle \nu, S \rangle_t + \frac{\partial^2 V_1}{\partial \nu^2} d\langle \nu \rangle_t \right] \right\} . \end{aligned}$$

Thereby we have for the quadratic (co-) variations using the system of equations (3.3.1):

$$\begin{aligned} d \langle S \rangle_t &= \nu S^2 dt \\ d \langle S, \nu \rangle_t &= d \langle \nu, S \rangle_t = \rho \sigma \beta \nu S dt \\ d \langle \nu \rangle_t &= \sigma^2 \beta^2 \nu dt. \end{aligned}$$

So we get:

$$\begin{aligned} d\Pi^\varphi &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial \nu} d\nu - \varphi_B r B dt - \varphi_S dS \\ &+ \frac{1}{2} \left[ \frac{\partial^2 V}{\partial S^2} \nu S^2 dt + \frac{\partial^2 V}{\partial \nu \partial S} \rho \sigma \beta \nu S dt + \frac{\partial^2 V}{\partial S \partial \nu} \rho \sigma \beta \nu S dt + \frac{\partial^2 V}{\partial \nu^2} \sigma^2 \beta^2 \nu dt \right] \\ &- \varphi_{V_1} \left\{ \frac{\partial V_1}{\partial t} dt + \frac{\partial V_1}{\partial S} dS + \frac{\partial V_1}{\partial \nu} d\nu \right. \\ &\quad \left. + \frac{1}{2} \left[ \frac{\partial^2 V_1}{\partial S^2} \nu S^2 dt + \frac{\partial^2 V_1}{\partial \nu \partial S} \rho \sigma \beta \nu S dt + \frac{\partial^2 V_1}{\partial S \partial \nu} \rho \sigma \beta \nu S dt + \frac{\partial^2 V_1}{\partial \nu^2} \sigma^2 \beta^2 \nu dt \right] \right\}. \end{aligned}$$

Collecting the  $dt$  terms, the  $dS$  terms and the  $d\nu$  terms and using the fact that  $\frac{\partial^2 V}{\partial \nu \partial S} = \frac{\partial^2 V}{\partial S \partial \nu}$  yields:

$$\begin{aligned} d\Pi^\varphi &= \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} \right\} dt - \varphi_B r B dt \\ &- \varphi_{V_1} \left\{ \frac{\partial V_1}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho \sigma \beta \nu S \frac{\partial^2 V_1}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V_1}{\partial \nu^2} \right\} dt \\ &+ \left\{ \frac{\partial V}{\partial S} - \varphi_{V_1} \frac{\partial V_1}{\partial S} - \varphi_S \right\} dS + \left\{ \frac{\partial V}{\partial \nu} - \varphi_{V_1} \frac{\partial V_1}{\partial \nu} \right\} d\nu. \end{aligned}$$

To make the portfolio instantaneously risk-free, we have to get rid of the  $dS$  term and the  $d\nu$  term, i.e. we have to require that

$$\frac{\partial V}{\partial S} - \varphi_{V_1} \frac{\partial V_1}{\partial S} - \varphi_S \stackrel{!}{=} 0 \quad \text{and} \quad \frac{\partial V}{\partial \nu} - \varphi_{V_1} \frac{\partial V_1}{\partial \nu} \stackrel{!}{=} 0. \quad (3.3.3)$$

This leaves us with

$$\begin{aligned} d\Pi^\varphi &= \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} \right\} dt - \varphi_B r B dt \\ &- \varphi_{V_1} \left\{ \frac{\partial V_1}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho \sigma \beta \nu S \frac{\partial^2 V_1}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V_1}{\partial \nu^2} \right\} dt. \end{aligned}$$

On the other hand we have if the portfolio is riskless by the no-arbitrage paradigm:

$$d\Pi^\varphi \stackrel{!}{=} r \Pi^\varphi dt = r(V - \varphi_B B - \varphi_S S - \varphi_{V_1} V_1) dt.$$

Since the money market account terms cancel out, we have:

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} \\ &- \varphi_{V_1} \left\{ \frac{\partial V_1}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho \sigma \beta \nu S \frac{\partial^2 V_1}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V_1}{\partial \nu^2} \right\} \\ &= r(V - \varphi_S S - \varphi_{V_1} V_1). \end{aligned}$$

From condition (3.3.3) we get:

$$\varphi_{V_1} = \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial V_1}{\partial \nu}} \quad \text{and} \quad \varphi_S = \frac{\partial V}{\partial S} - \frac{\frac{\partial V}{\partial \nu}}{\frac{\partial V_1}{\partial \nu}} \cdot \frac{\partial V_1}{\partial S}$$

Using this and collecting all  $V$  terms on the left-hand side and all  $V_1$  terms on the right-hand side then yields:

$$\begin{aligned} & \frac{\frac{\partial V}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma\beta\nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2}\sigma^2\beta^2\nu \frac{\partial^2 V}{\partial \nu^2} + rS \frac{\partial V}{\partial S} - rV}{\frac{\partial V}{\partial \nu}} \\ &= \frac{\frac{\partial V_1}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho\sigma\beta\nu S \frac{\partial^2 V_1}{\partial \nu \partial S} + \frac{1}{2}\sigma^2\beta^2\nu \frac{\partial^2 V_1}{\partial \nu^2} + rS \frac{\partial V_1}{\partial S} - rV_1}{\frac{\partial V_1}{\partial \nu}} \end{aligned} \quad (3.3.4)$$

Note that in equation (3.3.4) the left-hand side is a function of  $V$  only and the right-hand side is a function of  $V_1$  only. Since the options  $V$  and  $V_1$  will typically have different payoffs, strikes or expiries the only way that this equation holds is that both sides are independent of the contract type.  $\rightarrow$  Both sides have to be equal to some function  $f$  of the **independent** variables  $S, \nu$  and  $t$  and **not** of the options themselves.

We choose w.l.o.g.

$$f(S, \nu, t) := -(\alpha(S, \nu, t) - \Phi(S, \nu, t)\beta(S, \nu, t)) \quad ,$$

where we skip the dependences and write as short-hand notation

$$f(S, \nu, t) = -(\alpha - \Phi\beta) \quad .$$

$\Phi = \Phi(S, \nu, t)$  is called the **market price of volatility risk**, which will be explained below.

We finally get following pricing equation:

$$\begin{aligned} & \frac{\partial V}{\partial t} + \frac{1}{2}\nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma\beta\nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2}\sigma^2\beta^2\nu \frac{\partial^2 V}{\partial \nu^2} + rS \frac{\partial V}{\partial S} - rV \\ &= -(\alpha - \Phi\beta) \frac{\partial V}{\partial \nu} \quad . \end{aligned} \quad (3.3.5)$$

### The Market Price of Volatility Risk

No we explain why  $\Phi$  is called market price of volatility risk and why we have chosen  $f(S, \nu, t)$  as above. If we can solve equation (3.3.5) we have found the value of the option and the hedge ratios. But note that we then get *two* hedge ratios, namely  $\frac{\partial V}{\partial S}$  and  $\frac{\partial V}{\partial \nu}$ , since we have two sources of randomness. Because the volatility is not traded, the pricing equation should contain a market price of risk term.

Let us consider a self-financing trading strategy

$$\xi = (\xi_V = 1, -\xi_B, -\xi_S)$$

leading to a portfolio

$$\Pi^\xi = V - \xi_B B - \xi_S S$$

and suppose that the option corresponding to the price process  $V$  is already delta hedged (i.e. choose  $\xi_S = \frac{\partial V}{\partial S}$ ) and not vega hedged satisfying (3.3.5).

Again since  $\xi$  was assumed to be self-financing and by Ito's formula we get for the dynamics of the portfolio:

$$\begin{aligned}
d\Pi^\xi &= dV - d(\xi_B B) - d(\xi_S S) = dV - \xi_B dB - \xi_S dS \\
&= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial \nu} d\nu + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} d\langle S \rangle_t + \frac{1}{2} \frac{\partial^2 V}{\partial \nu^2} d\langle \nu \rangle_t + \frac{\partial^2 V}{\partial \nu \partial S} d\langle S, \nu \rangle_t \\
&\quad - \xi_B r B dt - \xi_S dS \\
&= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial \nu} d\nu + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} dt + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} dt + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} dt \\
&\quad - \xi_B r B dt - \xi_S dS \\
&= \left\{ \frac{\partial V}{\partial S} - \xi_S \right\} dS + \frac{\partial V}{\partial \nu} d\nu \\
&\quad + \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} - \xi_B r B \right\} dt \\
&\stackrel{(\diamond)}{=} \frac{\partial V}{\partial \nu} d\nu + \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} - \xi_B r B \right\} dt.
\end{aligned}$$

Thereby  $(\diamond)$  holds because the option is assumed to be delta hedged.

Now we look at the excess return<sup>4</sup> of the portfolio  $\Pi^\xi$ :

$$\begin{aligned}
d\Pi^\xi - r\Pi^\xi dt &= \frac{\partial V}{\partial \nu} d\nu + \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} - \xi_B r B \right\} dt \\
&\quad - r \{V - \xi_B B - \xi_S S\} dt \\
&= \frac{\partial V}{\partial \nu} d\nu + \underbrace{\left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \sigma^2 \beta^2 \nu \frac{\partial^2 V}{\partial \nu^2} + \rho \sigma \beta \nu S \frac{\partial^2 V}{\partial \nu \partial S} + rS \frac{\partial V}{\partial S} - rV \right\}}_{\stackrel{(3.3.5)}{=} -(\alpha - \Phi\beta) \frac{\partial V}{\partial \nu}} dt
\end{aligned}$$

So we get:

$$\begin{aligned}
d\Pi^\xi - r\Pi^\xi dt &= \frac{\partial V}{\partial \nu} d\nu - \frac{\partial V}{\partial \nu} (\alpha - \Phi\beta) dt \\
&\stackrel{(3.3.1)}{=} \frac{\partial V}{\partial \nu} (\alpha dt + \sigma \beta \sqrt{\nu} dW_2(t)) - \frac{\partial V}{\partial \nu} (\alpha - \Phi\beta) dt \\
&= \beta \sqrt{\nu} \frac{\partial V}{\partial \nu} \left\{ \Phi(S, \nu, t) dt + \frac{\sigma}{\sqrt{\nu}} dW_2(t) \right\}
\end{aligned}$$

We see - if we choose  $f(S, \nu, t) = -(\alpha - \Phi(S, \nu, t)\beta)$  - that for  $\frac{\sigma}{\sqrt{\nu}}$  units of volatility risk represented by  $dW_2(t)$ , there are  $\Phi$  units of extra return represented by  $dt$  (hence the name market price of volatility risk). That is why our choice of  $f(S, \nu, t)$  was reasonable.

We now get the pricing PDE corresponding to the Heston model if we substitute  $\alpha = \kappa(\theta - \nu)$  and  $\beta = 1$  in (3.3.5):

$$\frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sigma \nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 V}{\partial \nu^2} + rS \frac{\partial V}{\partial S} - rV = -(\kappa(\theta - \nu) - \Phi) \frac{\partial V}{\partial \nu}. \quad (3.3.6)$$

<sup>4</sup>By the excess return of a portfolio  $\Pi$  we understand what we get beyond the return of the portfolio  $\Pi$  if the portfolio  $\Pi$  were riskless.

Finally, we have to specify the form of the market price of volatility risk. Various economic arguments can be made (see for example [Heston, 1993]) that show that  $\Phi$  should be proportional to the variance, i. e. we choose

$$\Phi(s, \nu, t) = \lambda \nu. \quad (3.3.7)$$

for some constant  $\lambda$ . Another justification of the choice of  $\Phi$  is that we get a closed-form solution in the end which is a quite convincing argument. This finally yields the **pricing PDE corresponding the Heston model under the physical measure P**:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sigma \nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 V}{\partial \nu^2} + r S \frac{\partial V}{\partial S} - r V = -(\kappa(\theta - \nu) - \lambda \nu) \frac{\partial V}{\partial \nu}. \quad (3.3.8)$$

In order to get a solution for European calls we have to solve (3.3.8) using appropriate boundary conditions. This will be done in the next section.

**Remark1.** Defining the risk-adjusted parameters  $\kappa^*$  and  $\theta^*$  by

$$\kappa^* = \kappa + \lambda \quad \text{and} \quad \theta^* = \frac{\kappa \theta}{\kappa^*} = \frac{\kappa \theta}{\kappa + \lambda} \quad (3.3.9)$$

and substituting them in (3.3.8) we get the **risk-neutral pricing PDE corresponding to the Heston model**:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \nu S^2 \frac{\partial^2 V}{\partial S^2} + \rho \sigma \nu S \frac{\partial^2 V}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 V}{\partial \nu^2} + r S \frac{\partial V}{\partial S} - r V = -\kappa^*(\theta^* - \nu) \frac{\partial V}{\partial \nu}. \quad (3.3.10)$$

**Remark2.** Before solving the PDE (3.3.8), we have to pose us the following questions:

- Does a solution exist?
- If yes, is this solution unique?
- Can this solution be represented as a discounted expectation?

The answers to these questions are given in [Primm, 2007]. There it is shown under the use of [Korn & Korn, 2001] and [Heath & Schweizer, 2000] that one can extend the Feynman-Kac representation (compare how the solution of the Black-Scholes PDE in section 1.1 is derived) to the case of the pricing PDE according to the Heston model when European calls and puts are considered.

In particular, it is shown that the function  $u(t, (S, \nu)')$  is a *unique solution* of the PDE

$$\frac{1}{2} \nu S^2 \frac{\partial^2 u}{\partial S^2} + \rho \sigma \nu S \frac{\partial^2 u}{\partial \nu \partial S} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 u}{\partial \nu^2} + r S \frac{\partial u}{\partial S} + (\kappa(\theta - \nu) - \lambda \nu) \frac{\partial u}{\partial \nu} - r u + \frac{\partial u}{\partial t} = 0$$

subject to the initial condition

$$u(T, (S, \nu)') = \max \{K - S; 0\} \quad (\text{resp. } u(T, (S, \nu)') = \max \{S - K; 0\}).$$

Thereby the function  $u(t, (S, \nu)') \in C^{1,2}([0, T] \times D, \mathbb{R})$ , where  $D = (0, \infty)^2 \subseteq \mathbb{R}^2$ , has the representation

$$u(t, (S, \nu)') = \left[ h(X(T)) e^{-\int_t^T k(u, X(u)) du} + \int_t^T g(s, X(s)) e^{-\int_t^s k(u, X(u)) du} ds \right] \Big|_{X(t) = (S, \nu)'}$$



for given measurable functions  $h : D \rightarrow \mathbb{R}$ ,  $k : [0, T] \times D \rightarrow [0, \infty)$  and  $g : [0, T] \times D \rightarrow \mathbb{R}$ , where the 2-dimensional stochastic process  $\{X(t)\}_{0 \leq t \leq T}$  is defined by the SDE

$$dX(t) = b(t, X(t))dt + \sum_{j=1}^2 \Sigma_j(t, X(t))dW_j(t), \quad X(0) = x \in D$$

with the 2-dimensional Brownian motion  $W = (W_1, W_2)'$  and where the coefficients  $b : [0, T] \times D \rightarrow \mathbb{R}^2$  and  $\sigma : [0, T] \times D \rightarrow \mathbb{R}^{2 \times 2}$  corresponding to the risk-neutral processes<sup>5</sup> of the Heston model

$$\begin{aligned} dS(t) &= r(t)S(t)dt + \sqrt{\nu(t)}S(t)dW_1(t) \\ d\nu(t) &= \kappa^*(\theta^* - \nu(t))dt + \sigma\sqrt{\nu(t)}dW_2(t) \\ \text{Corr}(dW_1, dW_2) &= \rho \end{aligned} \quad (*)$$

are given by

$$b(t, (S, \nu)') = \begin{pmatrix} rS \\ \kappa(\theta - \nu) - \lambda\nu \end{pmatrix} \quad \text{and} \quad \Sigma(t, (S, \nu)') = \begin{pmatrix} \sqrt{\nu}S & 0 \\ \rho\sigma\sqrt{\nu} & \sqrt{1 - \rho^2}\sqrt{\nu} \end{pmatrix},$$

where we used that the SDEs (\*) with the correlated Brownian motions  $W_1$  and  $W_2$  can be written using two uncorrelated Brownian motions  $\widetilde{W}_1$  and  $\widetilde{W}_2$  as

$$\begin{aligned} dS(t) &= r(t)S(t)dt + \sqrt{\nu(t)}S(t)d\widetilde{W}_1(t) \\ d\nu(t) &= \kappa^*(\theta^* - \nu(t))dt + \sigma\sqrt{\nu(t)} \left\{ \rho d\widetilde{W}_1(t) + \sqrt{1 - \rho^2} d\widetilde{W}_2(t) \right\} \end{aligned}$$

(see [Heath & Schweizer, 2000]) and the fact that  $\kappa^*(\theta^* - \nu) = \kappa(\theta - \nu) - \lambda\nu$  by (3.3.9).

### 3.4 Closed-form Solution for European Calls

The above derivation of the pricing equation of the Heston model is only valid if there exists a  $C^{1,2}$ -solution. We now want to price a European call written on a stock  $S$  with strike  $K$ , maturity  $T$  and payoff  $C(T) = \max\{S(T) - K; 0\}$  with price process  $V = V(S, \nu, t) \in C^{1,2}$  such that  $C(t) = V(S(t), \nu(t), t)$ . We follow [Heston, 1993], [Gatheral, 2006] and [Primm, 2007]. A European call with corresponding price process  $V$  satisfies the PDE (3.3.8) subject to the following boundary conditions:

$$\begin{aligned} V(S, \nu, T) &= \max\{S - K; 0\}, \\ V(0, \nu, t) &= 0, \\ \frac{\partial V(\infty, \nu, t)}{\partial S} &= 1, \\ \frac{\partial V(S, 0, t)}{\partial t} + rS \frac{\partial V(S, 0, t)}{\partial S} - rV(S, 0, t) + \kappa\theta \frac{\partial V(S, 0, t)}{\partial \nu} &= 0, \\ V(S, \infty, t) &= S. \end{aligned} \quad (3.4.1)$$

So our goal is to solve the PDE (3.3.8) subject to the boundary conditions (3.4.1).

---

<sup>5</sup>We use here the risk-neutral form of the Heston processes, since solely these determine option prices by the no-arbitrage paradigm.

By analogy with the structure of the Black-Scholes formula (compare theorem 1.2) we use as a sophisticated guess for the solution

$$C(t) = C(S, \nu, t) = S \cdot P_1(S, \nu, t) - K \cdot p(t, T) \cdot P_2(S, \nu, t) \quad , \quad (3.4.2)$$

where the first term is the present value of the asset upon minimal exercise and the second term is the present value of the strike price payment. Both of these terms must satisfy the PDE (3.3.8).

### Digression 3: Change of Numéraire

During the course of this digression we will justify the form of our Ansatz (3.4.2). We follow [Kraft, 2006].

Under an equivalent martingale measure  $Q$  we can write the time- $t$  price of a European call as the discounted expectation of its terminal payoff w.r.t.  $Q$ :

$$\frac{C(t)}{B(t)} = {}^Q \left[ \frac{\max \{S(T) - K; 0\}}{B(T)} \right] = {}^Q \left[ \frac{S(T) \cdot 1_{\{S(T) \geq K\}}}{B(T)} \right] - {}^Q \left[ \frac{1_{\{S(T) \geq K\}}}{B(T)} \right] ,$$

where we discount with the money market account.

The choice of the discount factor as the the money market account is a posteriori quite arbitrary. Applying the change of numéraire technique we can use a discount factor different from the money market account. To do so, we give at first a basic definition.

#### Definition 3.4 (Numéraire Pair):

The pair  $(X, Q_X)$  is said to be a numéraire pair if

- (i)  $X$  is a price process, i.e. a portfolio of  $S_0, \dots, S_m$ .
- (ii)  $X > 0$ .
- (iii) The measure  $Q_X$  is equivalent to the physical measure  $P$ .
- (iv) The processes  $\left\{ \frac{S_j(t)}{X(t)} \right\}_{t \geq 0}$  are  $Q_X$ -martingales, i.e.  ${}^{Q_X} \left[ \frac{S_j(t)}{X(t)} \middle| \mathcal{F}_s \right] = \frac{S_j(s)}{X(s)} \quad \forall s \leq t$ .

**Remark.** The standard numéraire pair from basic lectures in financial mathematics is  $(B, Q)$ .

The following theorem how we can change the discounting factor - this is actually called change of numéraire - and how the option price is calculated w.r.t. a new discounting factor.

#### Theorem 3.5 (Change of Numéraire).

Let  $(X, Q_X)$  be a numéraire pair,  $Y > 0$  a price process and  $\left\{ \frac{Y(t)}{X(t)} \right\}_{t \geq 0}$  a  $Q_X$ -martingale. Then we have:

- (i) There exists a probability measure  $Q_Y$  defined by

$$\frac{dQ_Y}{dQ_X} \bigg|_{\mathcal{F}_t} = \frac{Y(t) \cdot X(0)}{X(t) \cdot Y(0)}$$

such that  $(Y, Q_Y)$  is a numéraire pair.

(ii) The time- $t$  price of a contingent claim with  $\frac{C(T)}{X(T)} \in L^1(Q_X)$  reads:

$$C(t) = X(t) \cdot {}^{Q_X} \left[ \frac{C(T)}{X(T)} \middle| \mathcal{F}_t \right] = Y(t) \cdot {}^{Q_Y} \left[ \frac{C(T)}{Y(T)} \middle| \mathcal{F}_t \right].$$

**Proof:**

See [Kraft, 2006]. ■

Assume now that  $\left\{ \frac{S(t)}{B(t)} \right\}_{t \geq 0}$  is a  $Q_S$ -martingale<sup>6</sup> and that  $\left\{ \frac{p(t,T)}{B(t)} \right\}_{t \geq 0}$  is a  $Q_P$ -martingale, then we can derive the Ansatz (3.4.2) applying the two changes of numéraire  $(B, Q) \rightarrow (S, Q^S)$  and  $(B, Q) \rightarrow (p(\cdot, T), Q^P)$ :

$$\begin{aligned} C(t) &= B(t) \cdot {}^Q \left[ \frac{\max\{S(T) - K; 0\}}{B(T)} \middle| \mathcal{F}_t \right] \\ &= B(t) \cdot {}^Q \left[ \frac{S(T)}{B(T)} \cdot 1_{\{S(T) > K\}} \middle| \mathcal{F}_t \right] - K \cdot B(t) \cdot {}^Q \left[ \frac{1}{B(T)} \cdot 1_{\{S(T) > K\}} \middle| \mathcal{F}_t \right] \\ &= S(t) \cdot {}^{Q_S} \left[ \frac{S(T)}{S(T)} \cdot 1_{\{S(T) > K\}} \middle| \mathcal{F}_t \right] - K \cdot p(t, T) \cdot {}^{Q_P} \left[ \frac{1}{p(T, T)} \cdot 1_{\{S(T) > K\}} \middle| \mathcal{F}_t \right] \\ &= S(t) \cdot {}^{Q_S} [1_{\{S(T) > K\}} | \mathcal{F}_t] - K \cdot p(t, T) \cdot {}^{Q_P} [1_{\{S(T) > K\}} | \mathcal{F}_t] \\ &= S(t) \cdot P_1(S(T) > K) - K \cdot p(t, T) \cdot P_2(S(T) > K) \end{aligned}$$

**Remarks.**

- This is a very general result and is valid for any equity price model.
- By changing the numeraires we can also fit the stock price process and the variance process to the new measures  $Q_S$  corresponding to the probability  $P_1$  and  $Q_P$  corresponding to the probability  $P_2$ .

(End of Digression 3)

Having motivated why the choice of the Ansatz (3.4.2) is reasonable we now finally want to solve (3.3.8). For solving it, it is convenient to use the transformation

$$x = \ln(S) \quad ,$$

since using this the dependence on  $S$  vanishes in (3.3.8).

We then get for the partial derivatives in the pricing equation:

$$\begin{aligned} \frac{\partial V}{\partial S} &= \frac{\partial V}{\partial x} \cdot \frac{\partial x}{\partial S} = \frac{\partial V}{\partial x} \cdot \frac{1}{S} \\ \frac{\partial^2 V}{\partial S^2} &= \frac{\partial}{\partial S} \left( \frac{1}{S} \frac{\partial V}{\partial x} \right) = -\frac{1}{S^2} \frac{\partial V}{\partial x} + \frac{1}{S} \frac{\partial}{\partial S} \left( \frac{\partial V}{\partial x} \right) = -\frac{1}{S^2} \frac{\partial V}{\partial x} + \frac{1}{S^2} \frac{\partial^2 V}{\partial x^2} \\ \frac{\partial^2 V}{\partial \nu \partial S} &= \frac{\partial}{\partial \nu} \left( \frac{\partial V}{\partial S} \right) = \frac{\partial}{\partial \nu} \left( \frac{1}{S} \frac{\partial V}{\partial x} \right) = \frac{1}{S} \frac{\partial^2 V}{\partial \nu \partial x} \end{aligned} \tag{3.4.3}$$

---

<sup>6</sup>The martingale property of  $\left\{ \frac{S(t)}{B(t)} \right\}_{t \geq 0}$  is for example shown in theorem A.3 of [Kruse & Nögel, 2005].

The partial derivatives w.r.t. the variance  $\frac{\partial V}{\partial \nu}$  and  $\frac{\partial^2 V}{\partial \nu^2}$  are invariant under the transformation. Using (3.4.3) we obtain from (3.3.8):

$$\begin{aligned} \frac{\partial V}{\partial t} - \frac{1}{2}\nu \frac{\partial V}{\partial x} + \frac{1}{2}\nu \frac{\partial^2 V}{\partial x^2} + \rho\sigma\nu \frac{\partial^2 V}{\partial \nu \partial x} + \frac{1}{2}\sigma^2\nu \frac{\partial^2 V}{\partial \nu^2} + r \frac{\partial V}{\partial x} - rV \\ + (\kappa(\theta - \nu) - \lambda\nu) \frac{\partial V}{\partial \nu} = 0 \end{aligned} \quad (3.4.4)$$

Due to the transformation  $x = \ln(S)$  the Ansatz (3.4.2) transforms to

$$C(S, \nu, t) = e^x \cdot P_1(x, \nu, t) - K \cdot p(t, T) \cdot P_2(x, \nu, t) \quad . \quad (3.4.5)$$

We get for the partial derivatives of the transformed Ansatz where we assume that the interest rate  $r$  is constant for the sake of simplicity:

$$\begin{aligned} \frac{\partial C}{\partial x} &= e^x \left( P_1 + \frac{\partial P_1}{\partial x} \right) - K p(t, T) \frac{\partial P_2}{\partial x} \\ \frac{\partial^2 C}{\partial x^2} &= e^x \left( P_1 + 2 \frac{\partial P_1}{\partial x} + \frac{\partial^2 P_1}{\partial x^2} \right) - K p(t, T) \frac{\partial^2 P_2}{\partial x^2} \\ \frac{\partial C}{\partial \nu} &= e^x \frac{\partial P_1}{\partial \nu} - K p(t, T) \frac{\partial P_2}{\partial \nu} \\ \frac{\partial^2 C}{\partial \nu^2} &= e^x \frac{\partial^2 P_1}{\partial \nu^2} - K p(t, T) \frac{\partial^2 P_2}{\partial \nu^2} \\ \frac{\partial^2 C}{\partial \nu \partial x} &= e^x \left( \frac{\partial P_1}{\partial \nu} + \frac{\partial^2 P_1}{\partial \nu \partial x} \right) - K p(t, T) \frac{\partial^2 P_2}{\partial \nu \partial x} \\ \frac{\partial C}{\partial t} &= e^x \frac{\partial P_1}{\partial t} - K \left( \frac{\partial p(t, T)}{\partial t} P_2 + p(t, T) \frac{\partial P_2}{\partial t} \right) = e^x \frac{\partial P_1}{\partial t} - K p(t, T) \left( r P_2 + \frac{\partial P_2}{\partial t} \right) \end{aligned} \quad (3.4.6)$$

Substituting (3.4.5) and (3.4.6) in (3.4.4) we get after collecting the  $P_1$  terms and the  $P_2$  terms:

$$\begin{aligned} e^x \left\{ \underbrace{\frac{\partial P_1}{\partial t} + \left( r + \frac{1}{2}\nu \right) \frac{\partial P_1}{\partial x} + \frac{1}{2}\nu \frac{\partial^2 P_1}{\partial x^2} + \rho\sigma\nu \frac{\partial^2 P_1}{\partial \nu \partial x} + \frac{1}{2}\sigma^2\nu \frac{\partial^2 P_1}{\partial \nu^2} + (\rho\sigma\nu + \kappa\theta - \kappa\nu - \lambda\nu) \frac{\partial P_1}{\partial \nu}}_{(*)} \right\} \\ - K p(t, T) \left\{ \underbrace{\frac{\partial P_2}{\partial t} + \left( r - \frac{1}{2}\nu \right) \frac{\partial P_2}{\partial x} + \frac{1}{2}\nu \frac{\partial^2 P_2}{\partial x^2} + \rho\sigma\nu \frac{\partial^2 P_2}{\partial \nu \partial x} + \frac{1}{2}\sigma^2\nu \frac{\partial^2 P_2}{\partial \nu^2} + (\kappa\theta - \kappa\nu - \lambda\nu) \frac{\partial P_2}{\partial \nu}}_{(**)} \right\} = 0 \end{aligned}$$

This equation can only hold if we have that  $(*) = 0$  and  $(**) = 0$ . By these conditions we get two partial differential equations for the probabilities  $P_1$  and  $P_2$ :

$$\frac{\partial P_j}{\partial t} + (r + u_j\nu) \frac{\partial P_j}{\partial x} + \frac{1}{2}\nu \frac{\partial^2 P_j}{\partial x^2} + \rho\sigma\nu \frac{\partial^2 P_j}{\partial \nu \partial x} + \frac{1}{2}\sigma^2\nu \frac{\partial^2 P_j}{\partial \nu^2} + (a - b_j\nu) \frac{\partial P_j}{\partial \nu} = 0, \quad (3.4.7)$$

for  $j = 1, 2$ , where

$$u_1 = \frac{1}{2}, \quad u_2 = -\frac{1}{2}, \quad a = \kappa\theta, \quad b_1 = \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda. \quad (3.4.8)$$

For the option price to satisfy the terminal condition  $C(T) = C(S, \nu, T) = \max\{S - K; 0\}$  of a European call the probabilities  $\{P_j\}_{j=1,2}$  have to fulfill the terminal condition

$$P_j(x, \nu, T; \ln[K]) = 1_{\{x \geq \ln[K]\}}. \quad (3.4.9)$$

**Proposition 3.6.**

Let the stochastic processes  $x(t)$  for  $j = 1, 2$  be given by<sup>7</sup>

$$\begin{aligned} dx(t) &= (r + u_j \nu(t))dt + \sqrt{\nu(t)}dW_1(t) \\ d\nu(t) &= (a_j - b_j \nu(t))dt + \sigma \sqrt{\nu(t)}dW_2(t) \end{aligned} \quad (3.4.10)$$

with  $\text{Corr}(dW_1, dW_2) = \rho dt$ , where the parameters  $u_j$ ,  $a_j$  and  $b_j$  are defined by (3.4.8). Then  $P_j$  is the conditional probability that the option expires in-the-money, i.e.

$$P_j(x, \nu, t; \ln[K]) = P(x(T) \geq \ln[K] | x(t) = x, \nu(t) = \nu) \quad j = 1, 2.$$

Additionally, the characteristic functions  $\varphi_{t,x(T)}^1(u)$  and  $\varphi_{t,x(T)}^2(u)$  corresponding to  $P_1$  and  $P_2$  fulfill the PDEs (3.4.7) w.r.t. the terminal condition

$$\varphi_{T,x(T)}^j(u) = e^{iux}.$$

**Proof:**

Let  $f(x, \nu, t)$  be a twice differentiable function that is the conditional expectation of some function  $g(x, \nu)$  at a later date  $T$ , i.e.

$$f(x, \nu, t) = [g(x(T), \nu(T)) | x(t) = x, \nu(t) = \nu]. \quad (*)$$

By iterated expectations we get  $\forall s \geq t$ :

$$\begin{aligned} [f(x, \nu, s) | \mathcal{F}_t] &= [ [g(x(T), \nu(T)) | x(s) = x, \nu(s) = \nu] | \mathcal{F}_t] \\ &= [ [g(x(T), \nu(T)) | \mathcal{F}_s] | \mathcal{F}_t] \\ &= [g(x(T), \nu(T)) | \mathcal{F}_s], \text{ since } \mathcal{F}_t \subseteq \mathcal{F}_s \\ &= [g(x(T), \nu(T)) | x(t) = x, \nu(t) = \nu] \\ &= f(x, \nu, t). \end{aligned}$$

Therefore  $f$  is a martingale.

On the other hand, Ito's formula yields:

$$df = \frac{\partial f}{\partial t} dt dx + \frac{\partial f}{\partial \nu} d\nu + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} d \langle x \rangle_t + \frac{\partial^2 f}{\partial x \partial \nu} d \langle x, \nu \rangle_t + \frac{1}{2} \frac{\partial^2 f}{\partial \nu^2} d \langle \nu \rangle_t.$$

With  $d \langle x \rangle_t = \nu dt$ ,  $d \langle x, \nu \rangle_t = \rho \sigma \nu dt$  and  $d \langle \nu \rangle_t = \sigma^2 \nu dt$  and using the dynamics (3.4.10) we then get:

$$\begin{aligned} df &= \left\{ \frac{1}{2} \nu \frac{\partial^2 f}{\partial x^2} + \rho \sigma \nu \frac{\partial^2 f}{\partial x \partial \nu} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 f}{\partial \nu^2} + (r + u_j \nu) \frac{\partial f}{\partial x} + (a - b_j \nu) \frac{\partial f}{\partial \nu} + \frac{\partial f}{\partial t} \right\} dt \\ &\quad + \sqrt{\nu} \frac{\partial f}{\partial x} dW_1(t) + \sigma \sqrt{\nu} dW_2(t) \end{aligned}$$

Since we know that  $f$  is a martingale we can conclude by the martingale representation theorem that the  $dt$  term vanishes, i.e.

$$\frac{1}{2} \nu \frac{\partial^2 f}{\partial x^2} + \rho \sigma \nu \frac{\partial^2 f}{\partial x \partial \nu} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 f}{\partial \nu^2} + (r + u_j \nu) \frac{\partial f}{\partial x} + (a - b_j \nu) \frac{\partial f}{\partial \nu} + \frac{\partial f}{\partial t} = 0. \quad (**)$$

<sup>7</sup>Note that the processes  $x(t)$  and  $\nu(t)$  are different for the cases  $j = 1, 2$  just as the function  $f$  and we omit these indices for the sake of readability. Further, the Brownian motions are  $Q_S$ -Brownian motions ( $j = 1$ ) and  $Q_P$ -Brownian motions ( $j = 2$ ).

Equation (\*) thereby fixes the terminal condition:  $f(s, \nu, T) = g(x, \nu)$ .

Equation (\*\*) is the backward Kolmogorov equation for the system of SDEs (3.4.10). To see this, we cite from [Wilmott, 2006]:

Consider the SDE  $dy = A(y, t)dt + B(y, t)dW(t)$  and its corresponding transition probability density function  $\varphi = \varphi(y, t; y', t')$  with  $t' > t$  (remember section 2.1, where we derived the Fokker-Planck equation). Then  $\varphi$  fullfills the backward Kolmogorov equation

$$\frac{\partial \varphi}{\partial t} + \frac{1}{2}B(y, t)^2 \frac{\partial^2 \varphi}{\partial y^2} + A(y, t) \frac{\partial \varphi}{\partial y} = 0.$$

By analogy with this equation we then see that (\*\*) is indeed the backward Kolmogorov equation corresponding to (3.4.10). This backward equation is used to calculate probabilities of reaching a specific final state from various initial states. To be able to solve it, we need a terminal condition at a future state  $T > t$ .

If we now choose  $g(x, \nu) = 1_{\{x \geq \ln[K]\}}$  as terminal condition for equation (\*\*) at time  $T$  and since we know that (\*\*) exactly corresponds to the equations (3.4.7) derived for the probabilities  $P_1$  and  $P_2$ , we see that the corresponding solution  $f(S, \nu, t)$  at time  $t$  is the conditional probability at time  $t$  that  $x(T)$  is greater than  $\ln[K]$ . In other words:

$$f(S, \nu, t) = P_j(x, \nu, t; \ln[K]) = P(x(T) \geq \ln[K] | x(t) = x, \nu(t) = \nu) \quad j = 1, 2$$

is the conditinal probability that the option expires in the money. By this the first part of the proposition is proved.

It remains to show that the characteristic functions

$$\varphi_{x(T)}^j(x, \nu, t; u) = \left[ e^{iu x(T)} | x(t) = x, \nu(t) = \nu \right]$$

fullfill the PDEs (3.4.7).

Therefore just choose  $g(x, \nu) = e^{iu x}$  and the result follows. ■

Since the probabilities  $P_1$  and  $P_2$  are not immediately available in closed-form, we do not solve the equations (3.4.7) for  $P_1$  and  $P_2$ , i.e. w.r.t. the terminal conditions

$$g(x, \nu) = 1_{\{x \geq \ln[K]\}} = P_j(x, \nu, T; \ln[K]) \quad j = 1, 2.$$

Instead, we solve them w.r.t. the terminal conditions

$$g(x, \nu) = e^{iu x} = \varphi_{x(T)}^j(x, \nu, T; u) \quad j = 1, 2 \quad (3.4.11)$$

to obtain the characteristic functions.

At first, we invert the time direction by using the substitution

$$\tau := T - t.$$

Then the partial derivative w.r.t. in (3.4.7) transforms to

$$\frac{\partial P_j}{\partial t} = \frac{\partial P_j}{\partial \tau} \underbrace{\frac{\partial \tau}{\partial t}}_{=-1} = -\frac{\partial P_j}{\partial \tau}$$

and therefore (3.4.7) reads:

$$-\frac{\partial P_j}{\partial \tau} + (r + u_j \nu) \frac{\partial P_j}{\partial x} + \frac{1}{2} \nu \frac{\partial^2 P_j}{\partial x^2} + \rho \sigma \nu \frac{\partial^2 P_j}{\partial \nu \partial x} + \frac{1}{2} \sigma^2 \nu \frac{\partial^2 P_j}{\partial \nu^2} + (a - b_j \nu) \frac{\partial P_j}{\partial \nu} = 0. \quad (3.4.12)$$

Due to the linearity of the coefficients we use as a sophisticated guess for the solution

$$\varphi_{x(T)}^j(x, \nu, t; u) = e^{C_j(\tau; u) + D_j(\tau; u) \cdot \nu + i u x} \quad j = 1, 2. \quad (3.4.13)$$

We get for the partial derivatives of this functional form:

$$\begin{aligned} \frac{\partial \varphi_j}{\partial \tau} &= \left( \frac{\partial C_j}{\partial \tau} + \nu \frac{\partial D_j}{\partial \tau} \right) \varphi_j \\ \frac{\partial \varphi_j}{\partial x} &= i u \varphi_j \\ \frac{\partial^2 \varphi_j}{\partial x^2} &= -u^2 \varphi_j \\ \frac{\partial \varphi_j}{\partial \nu} &= D_j \varphi_j \\ \frac{\partial^2 \varphi_j}{\partial \nu^2} &= D_j^2 \varphi_j \\ \frac{\partial \varphi_j}{\partial \nu \partial x} &= i u D_j \varphi_j. \end{aligned} \quad (3.4.14)$$

By substituting the functional form (3.4.13) and the corresponding derivatives (3.4.14) in (3.4.12) we get after collecting terms:

$$\varphi_j \left\{ \underbrace{-\frac{\partial C_j}{\partial \tau} + a D_j + r u i}_{(*)} \right\} + \varphi_j \nu \left\{ \underbrace{-\frac{\partial D_j}{\partial \tau} + u_j u i - \frac{1}{2} u^2 + \rho \sigma u i D_j + \frac{1}{2} \sigma^2 D_j^2 - b_j D_j}_{(**)} \right\} = 0.$$

Since this equation can only hold if  $(*) = 0$  and  $(**) = 0$ , we can reduce the PDEs (3.4.7) to a set of two *coupled ordinary differential equations* for each  $j = 1, 2$ :

$$-\frac{1}{2} u^2 + \rho \sigma u i D_j + \frac{1}{2} \sigma^2 D_j^2 + u_j u i - b_j D_j - \frac{\partial D_j}{\partial \tau} = 0 \quad (3.4.15)$$

$$r u i + a D_j - \frac{\partial C_j}{\partial \tau} = 0 \quad (3.4.16)$$

subject to the initial conditions

$$C_j(0; u) = 0 \quad ; \quad D_j(0; u) = 0. \quad (3.4.17)$$

We directly see by substituting these initial conditions in the Ansatz (3.4.13) and keeping in mind that  $\tau = 0 \Leftrightarrow t = T$  that the terminal condition (3.4.11) is fulfilled.

Equation (3.4.15) is a *Ricatti differential equation*. We use the substitution

$$D_j(\tau; u) = -\frac{\frac{\partial E_j(\tau; u)}{\partial \tau}}{\frac{\sigma^2}{2} E_j(\tau; u)} \quad (3.4.18)$$

and obtain for (3.4.15):

$$\frac{\partial^2 E_j}{\partial \tau^2} - (\rho u i - b_j) \frac{\partial E_j}{\partial \tau} + \frac{\sigma^2}{2} \left( -\frac{1}{2} u^2 + u_j u i \right) E_j = 0.$$

These equations are *linear second order ordinary differential equations* and have the general solution

$$E_j(\tau; u) = A_j e^{x_{j,+} \cdot \tau} + B_j e^{x_{j,-} \cdot \tau}$$

with

$$x_{j,\pm} = \frac{\rho\sigma ui - b_j \pm d_j}{2} \quad \text{and} \quad d_j = \sqrt{(\rho\sigma ui - b_j)^2 - \sigma^2(2u_j ui - u^2)}.$$

Note thereby that

$$x_{j,+} - x_{j,-} = d_j \quad (3.4.19)$$

and define

$$\frac{x_{j,-}}{x_{j,+}} = \frac{b_j - \rho\sigma ui + d_j}{b_j - \rho\sigma ui - d_j} =: g_j. \quad (3.4.20)$$

From (3.4.17) and the substitution (3.4.18) we get for the initial conditions of  $E_j$ :

$$E_j(0; u) = A_j + B_j \quad \text{and} \quad \left. \frac{\partial E_j(\tau; u)}{\partial \tau} \right|_{\tau=0} = x_{j,+}A_j + x_{j,-}B_j = 0$$

These yield using the relation (3.4.20) the coefficients  $A_j$  and  $B_j$ :

$$A_j = \frac{g_j E_j(0; u)}{g_j - 1} \quad \text{and} \quad B_j = -\frac{E_j(0; u)}{g_j - 1}.$$

Thus we get

$$E_j(\tau; u) = \frac{E_j(0; u)}{g_j - 1} (g_j e^{x_{j,+}\cdot\tau} - e^{x_{j,-}\cdot\tau}) \quad \text{and} \quad \frac{\partial E_j(\tau; u)}{\partial \tau} = \frac{E_j(0; u)}{g_j - 1} (g_j x_{j,+} e^{x_{j,+}\cdot\tau} - x_{j,-} e^{x_{j,-}\cdot\tau}).$$

This finally yields for  $D_j$  using the relations (3.4.19), (3.4.20) and the definition of  $x_{j,-}$ :

$$\begin{aligned} D_j(\tau; u) &= -\frac{2}{\sigma^2} \frac{\frac{\partial E_j(\tau; u)}{\partial \tau}}{E_j(\tau; u)} = -\frac{2}{\sigma^2} \frac{g_j x_{j,+} e^{x_{j,+}\cdot\tau} - x_{j,-} e^{x_{j,-}\cdot\tau}}{g_j e^{x_{j,+}\cdot\tau} - e^{x_{j,-}\cdot\tau}} \\ &= -\frac{2}{\sigma^2} x_{j,-} \frac{e^{x_{j,+}\cdot\tau} - e^{x_{j,-}\cdot\tau}}{g_j e^{x_{j,+}\cdot\tau} - e^{x_{j,-}\cdot\tau}} = -\frac{2}{\sigma^2} x_{j,-} \frac{e^{x_{j,-}\cdot\tau} - e^{x_{j,+}\cdot\tau}}{e^{x_{j,-}\cdot\tau} - g_j e^{x_{j,+}\cdot\tau}} \\ &= -\frac{2}{\sigma^2} x_{j,-} \frac{1 - e^{(x_{j,+} - x_{j,-})\cdot\tau}}{1 - g_j e^{(x_{j,+} - x_{j,-})\cdot\tau}} = \frac{b_j - \rho\sigma ui + d_j}{\sigma^2} \cdot \left[ \frac{1 - e^{d_j\cdot\tau}}{1 - g_j e^{d_j\cdot\tau}} \right]. \end{aligned}$$

Having obtained the solution of (3.4.15), we can solve (3.4.16) by mere integration:

$$\begin{aligned} C_j(\tau; u) &= \int_0^\tau (rui + aD_j(\eta; u)) d\eta \\ &= rui\tau - \frac{2a}{\sigma^2} \int_0^\tau \frac{\frac{\partial E_j(\eta; u)}{\partial \tau}}{E_j(\eta; u)} d\eta = rui\tau - \frac{2a}{\sigma^2} \ln \left( \frac{E(\tau; u)}{E(0; u)} \right) \\ &= rui\tau - \frac{2a}{\sigma^2} \ln \left[ \frac{g_j e^{x_{j,+}\cdot\tau} - e^{x_{j,-}\cdot\tau}}{g_j - 1} \right] = rui\tau - \frac{2a}{\sigma^2} \ln \left[ \frac{e^{x_{j,-}\cdot\tau} g_j - e^{x_{j,+}\cdot\tau}}{1 - g_j} \right] \\ &= rui\tau - \frac{2a}{\sigma^2} \ln \left[ \frac{1 - g_j e^{(x_{j,+} - x_{j,-})\cdot\tau}}{e^{-x_{j,-}\cdot\tau} (1 - g_j)} \right] = rui\tau - \frac{2a}{\sigma^2} \ln \left[ x_{j,-}\tau + \frac{1 - g_j e^{d_j\cdot\tau}}{1 - g_j} \right] \\ &= rui\tau - \frac{a}{\sigma^2} \left\{ (\rho\sigma ui - b_j - d_j)\tau + 2 \ln \left[ \frac{1 - g_j e^{d_j\cdot\tau}}{1 - g_j} \right] \right\} \\ &= rui\tau + \frac{a}{\sigma^2} \left\{ (b_j - \rho\sigma ui + d_j)\tau - 2 \ln \left[ \frac{1 - g_j e^{d_j\cdot\tau}}{1 - g_j} \right] \right\} \end{aligned}$$



Thereby we used again the relations (3.4.19), (3.4.20) and the definition of  $x_{j,-}$ . Altogether we obtain for the solution of the coupled ODEs (3.4.15) and (3.4.16):

$$\begin{aligned} C_j(\tau; u) &= rui\tau + \frac{a}{\sigma^2} \left\{ (b_j - \rho\sigma ui + d_j)\tau - 2 \ln \left[ \frac{1 - g_j e^{d_j \tau}}{1 - g_j} \right] \right\}, \\ D_j(\tau; u) &= \frac{b_j - \rho\sigma ui + d_j}{\sigma^2} \cdot \left[ \frac{1 - e^{d_j \tau}}{1 - g_j e^{d_j \tau}} \right], \end{aligned} \quad (3.4.21)$$

where

$$\begin{aligned} g_j &= \frac{b_j - \rho\sigma ui + d_j}{b_j - \rho\sigma ui - d_j}, \\ d_j &= \sqrt{(\rho\sigma ui - b_j)^2 - \sigma^2(2u_j ui - u^2)} \end{aligned} \quad (3.4.22)$$

and  $u_j, b_j$  and  $a$  are already defined by (3.4.8) and  $\tau := T - t$ . So the characteristic functions  $\varphi_{x(T)}^j(x, \nu, t; u)$ ,  $j = 1, 2$  are completely determined by the equations (3.4.13), (3.4.21) and (3.4.22).

What is left now, is to get from the characteristic functions of the probabilities  $P_1$  and  $P_2$  to the probabilities themselves. When this is done, we have finally derived a close-form solution for European calls. The following digression will show how characteristic functions and the corresponding distribution functions are connected in general. Then we apply the results to the Heston case.

#### Digression 4: Characteristic Functions and Distribution Functions

During the course of this digression we follow [Jacod & Protter, 2004], [Dufresne et al., 2005], [Korn, 2005] and [Primm, 2007].

**Notation:** Let  $X$  be a real-valued random variable. We then denote  $\mu_X$  the measure on induced by  $X$  given by  $\mu_X(B) = P\{X \in B\}$ , where  $B$  is a Borel subset of  $\mathbb{R}$  and  $F_X$  the distribution function of  $X$  given by  $F_X(x) = P\{X \leq x\}$ ,  $x \in \mathbb{R}$ .

We already used the notion of a characteristic function above. At first, we define what we mean by a characteristic function corresponding to a random variable  $X$  and give some properties of characteristic functions.

#### Definition 3.7 (Characteristic Function):

Let  $X$  be a real-valued random variable and let  $\mu_X$  be the corresponding probability measure. Then its characteristic function  $\varphi_X$  is defined as

$$\varphi_X(u) := [e^{iuX}] = \int_{-\infty}^{+\infty} e^{iux} \mu_X(dx).$$

#### Lemma 3.8 (Properties of characteristic Functions):

Let  $\varphi_X$  be the characteristic function of a real-valued random variable. Then we have:

- (i)  $\varphi_X$  always exists, i.e. it is well-defined.
- (ii)  $\varphi_X$  is a bounded function.
- (iii)  $\varphi_X(0) = 1$ .
- (iv)  $\varphi_X(-u) = \overline{\varphi_X(u)}$ .

**Proof:**

ad(i):

$\varphi_X$  always exists if the expectation always exists. We have that

$$\varphi_X(u) := [e^{iuX}] = [\cos(uX)] + i [\sin(uX)] < \infty ,$$

since sine and cosine are bounded functions. So (i) is proved.

ad(ii):

$$|\varphi_X(u)| \leq \int_{-\infty}^{\infty} |e^{iuX}| \mu(dx) = \int_{-\infty}^{\infty} 1 \mu(dx) = 1 .$$

ad(iii):

$$\varphi_X(0) = \int_{-\infty}^{\infty} 1 \mu(dx) = 1 .$$

ad(iv):

$$\varphi_X(-u) = [e^{-iuX}] = [\cos(-uX)] + i [\sin(-uX)] = [\cos(uX)] - i [\sin(uX)] = \overline{\varphi_X(u)} .$$

■

Characteristic functions are connected with Fourier transforms. To see this, we give the following definition:

**Definition 3.9 (Fourier Transforms):**

For  $1 \leq p \leq \infty$  let  $L^p$  denote the space of measurable functions  $f : \mathbb{R} \rightarrow \mathbb{C}$  such that  $\int_{-\infty}^{\infty} |f(x)|^p dx < \infty$ .

- Let  $\mu$  be a signed measure on  $(\mathcal{B}(\mathbb{R}), \mathbb{R})$  with  $|\mu| < \infty$ . Then its Fourier transform is defined as

$$\hat{\mu}(u) = \int_{-\infty}^{\infty} e^{iux} \mu(dx) .$$

Since  $|e^{iux}| = 1$ , this is an ordinary Lebesgue integral.

- Let  $f \in L^1$ , then its Fourier transform  $\hat{f}$  is defined as

$$\hat{f}(u) = \int_{-\infty}^{\infty} e^{iux} f(x) dx .$$

- Let  $\hat{f} \in L^1$  be the Fourier transform of  $f \in L^1$ , then its inverse Fourier transform  $f$  is defined as

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \hat{f}(u) du .$$

**Remark.** Note that the requirement  $\hat{f} \in L^1$  (otherwise the Lebesgue integral of the inverse Fourier transform would not exist) is quite restrictive, since for a function  $f \in L^1$  not necessarily  $\hat{f} \in L^1$ . For example the Fourier transform of the exponential density function is not in  $L^1$ .

If we remember the definition of the characteristic function of a random variable, we see that it is the Fourier transform of the probability measure  $\mu_X$ :

Let  $X$  be a real-valued random variable and  $\mu_X$  the corresponding probability measure, then the Fourier transform  $\widehat{\mu}_X$  is given by

$$\widehat{\mu}_X(u) = \int_{-\infty}^{\infty} e^{iux} \mu_X(dx) = \varphi_X(u).$$

and therefore equal to the characteristic function  $\varphi_X$ .

Now we are able to connect the characteristic function  $\varphi_X$  with the distribution function  $F_X$  of  $X$ , which is given by

$$F_X(x) = P(X \leq x).$$

Does the random variable  $X$  now admit a density  $p_X$ , we can write:

$$\varphi_X(u) = \int_{-\infty}^{\infty} e^{iux} p_X(x) dx.$$

So if we know the characteristic function  $\varphi_X$  we can extract the density  $p_X$  by applying the inverse Fourier transformation:

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \varphi_X(u) du.$$

Integration of the density then yields the distribution function  $F_X$ :

$$F_X(x) = \int_{-\infty}^{\infty} p_X(s) ds = \int_{-\infty}^{\infty} \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ius} \varphi_X(u) du \right) ds. \quad (3.4.23)$$

**Remark.** We have shown by equation (3.4.23) how characteristic functions and distribution functions are connected in theory. However, formula (3.4.23) is of limited practical value, since on the one hand not every random variable admits a density function and on the other hand the calculation of the inverse Fourier transform of the characteristic function and the subsequent integration can be very complicated. Further as already mentioned above the Lebesgue-integral of the inverse Fourier transform does not necessarily have to exist. So we are interested in finding an inversion formula for the case of a general probability distribution  $\mu_X$  by which we can extract the distribution function  $F_X$  from the knowledge of the characteristic function  $\varphi_X$ .

Another problem is that the expectation  $[g(X)]$  is only defined for functions  $g \in L^1$ , but the functions we are interested in like  $g(x) = \max\{e^x - K; 0\}$  (compare (3.4.9)) are *not* in  $L^1$ . To solve this problem, one can use an exponential damping factor. Therefore we define for an arbitrary function  $f$  the function  $f^\alpha$  by

$$f^\alpha(x) = e^{\alpha x} f(x), \quad x \in \mathbb{R}.$$

The Fourier transform of  $f^\alpha$  is then given by:

$$\widehat{f^\alpha}(u) = \int_{-\infty}^{\infty} e^{iux} f^\alpha(x) dx = \int_{-\infty}^{\infty} e^{i(u-i\alpha)x} f(x) dx = \widehat{f}(u - i\alpha).$$

If we choose  $\alpha$  appropriate, the damped function  $g^{-\alpha}$  is in  $L^1$ .

This idea can also be applied to probability measures, in particular to  $\mu_X$ . For  $\alpha \in \mathbb{R}$  define a new probability measure  $\mu_X^\alpha$  by

$$\mu_X^\alpha(dx) := e^{\alpha x} \mu_X(dx).$$

Again we get for the Fourier transform of  $\mu_X^\alpha$ :

$$\widehat{\mu_X^\alpha}(u) = \int_{-\infty}^{\infty} e^{iux} \mu_X^\alpha(dx) = \int_{-\infty}^{\infty} e^{i(u-i\alpha)x} \mu_X(dx) = \widehat{\mu_X}(u - i\alpha).$$

To finally be able to prove a general inversion formula we still need the following lemma:

**Lemma 3.10:**

Let  $X$  be a real-valued random variable with corresponding probability measure  $\mu_X$  and suppose that for a particular  $\alpha \in \mathbb{R}$  the following conditions are fulfilled:

- (a)  $[e^{\alpha X}] < \infty$ ,
- (b)  $g^{-\alpha} \in L^1$ ,
- (c) the function  $y \mapsto [g(y + X)]$  is continuous at the origin and
- (d)  $f$  either satisfies condition (d1) or d(2):

(d1)  $f(y^+)$  and  $f(y^-)$  both exist and the integrals below are finite for some  $\epsilon > 0$ :

$$\int_0^\epsilon \frac{f(y+t) - f(y^+)}{t} dt, \quad \int_{-\epsilon}^0 \frac{f(y-t) - f(y^-)}{t} dt.$$

(d2)  $f(y)$  has bounded variation in some open neighbourhood of  $y$ .

Then we have:

$$[g(X)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{g^{-\alpha}}(-u) \widehat{\mu_X^\alpha}(u) du = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{g}(-u + i\alpha) \widehat{\mu_X}(u - i\alpha) du.$$

**Proof:**

We refer to the proof of theorem 2.1 of [Dufresne et al., 2005]. ■

Now we have provided everything we need to prove a general inversion formula which gives us the distribution function if the characteristic function is known. We get the following theorem:

**Theorem 3.11 (General Inversion Formula).**

Let  $F_X$  the distribution function corresponding to a real-valued random variable  $X$  be continuous at the point  $x = b$ , then we have:

$$F_X(b) = \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} [e^{iub} \widehat{\mu_X}(-u) - e^{-iub} \widehat{\mu_X}(u)] du \quad (3.4.24)$$

where the Fourier transform  $\widehat{\mu_X}$  of the probability measure  $\mu_X$  is equal to the characteristic function  $\varphi_X$  of the random variable  $X$ .

**Proof:**

If we substitute  $u \rightarrow -u$  in (3.4.24), we see that this formula is equivalent to

$$1 - F_X(b) = \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} [e^{-iub} \widehat{\mu_X}(u) - e^{+iub} \widehat{\mu_X}(-u)] du.$$

At first, we prove this formula for the case  $b = 0$ ; the general case is then obtained by the translation  $X \rightarrow X - b$ . For  $b = 0$  we have:

$$1 - F_X(0) = P(X \geq 0) = [1_{(0,\infty)}(X)].$$

$\Rightarrow$  We have to choose  $g(x) = 1_{(0,\infty)}(x)$  in lemma 3.10. However,  $1_{(0,\infty)}(x) \notin L^1$  and for  $u \in$  the Fourier transform does not exist since we have:

$$\widehat{1_{(0,\infty)}}(u) = \lim_{M \rightarrow \infty} \int_0^M e^{iux} dx = \lim_{M \rightarrow \infty} \left[ \frac{1}{iu} e^{iux} \right]_0^M = \lim_{M \rightarrow \infty} \frac{1}{iu} (e^{iuM} - 1) = \infty.$$

But for  $z = p + iq \in$  with  $Im(z) = q > 0$  the Fourier transform  $\widehat{1_{(0,\infty)}}(z)$  exists, since we have:

$$\widehat{1_{(0,\infty)}}(z) = \lim_{M \rightarrow \infty} \int_0^M e^{izx} dx = \lim_{M \rightarrow \infty} \frac{1}{iz} (e^{izM} - 1) = \lim_{M \rightarrow \infty} \frac{1}{iz} \underbrace{(e^{-qx})}_{\rightarrow 0} \underbrace{e^{ipM}}_{\text{bounded}} - 1 = -\frac{1}{iz}.$$

For the moment assume that there exists an  $\alpha > 0 \in$  such that  $[e^{\alpha X}] < \infty$  (this will be proved at the end of this proof), then we get by lemma 3.10:

$$1 - F_X(0) = [1_{(0,\infty)}(X)] = \frac{1}{2\pi} \int_{-\infty}^\infty \widehat{1_{(0,\infty)}}(-u + i\alpha) \widehat{\mu_X}(u - i\alpha) du. \quad (*)$$

Now define the function  $h$  as

$$h(z) := \widehat{1_{(0,\infty)}}(-z) \widehat{\mu_X}(z) = \frac{1}{iz} \widehat{\mu_X}(z) = \frac{[e^{izX}]}{iz}.$$

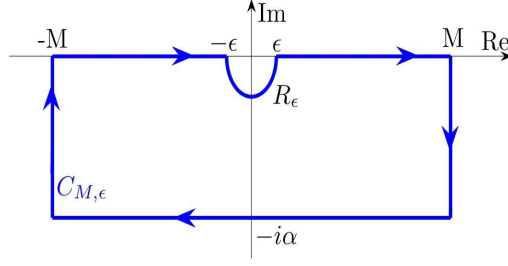
$h$  is analytic in the domain  $\{z \in \mathbb{C} \mid -\alpha < \Im m(z) < 0\}$  and has a pole at  $z = 0$ . Now we integrate  $h$  and choose the closed path  $C_{M,\epsilon}$  for integration (compare figure 10). So we can apply the Cauchy theorem of complex analysis which states that the integral of an in a domain analytic function over a closed path (which lies itself in this domain) is equal to zero. Using this we get:

$$\int_{C_{M,\epsilon}} h(z) dz = 0.$$

Now we decompose the path  $C_{M,\epsilon}$  in 4 subpaths and calculate the values of their corresponding integrals separately:

$$\int_{C_{M,\epsilon}} h(z) dz = \int_{L_{M,\epsilon}} h(z) dz + \int_{(1)}^{M-i\alpha} h(z) dz + \int_{(2)}^{-M-i\alpha} h(z) dz + \int_{(3)}^{-M} h(z) dz = 0. \quad (**)$$

Thereby  $L_{M,\epsilon}$  is the subpath going along the real axis from  $-M$  to  $-\epsilon$ , then around the half-circle  $R_\epsilon$  and then on the real axis from  $\epsilon$  to  $M$  (compare figure 10). For integral (3) we

Figure 10: Integration Path  $C_{M,\epsilon}$ 

have:

$$\int_{M-i\alpha}^{-M-i\alpha} h(z) dz = - \int_{-M-i\alpha}^{M-i\alpha} h(z) dz.$$

Thereby the imaginary part  $\Im m(z) = -\alpha$  of the integration variable  $z$  is fixed and the real part  $\Re e(z)$  runs from  $-M$  to  $M$ . This yields then by taking the limit  $M \rightarrow \infty$  for (3) using the definition of the function  $h$ :

$$- \int_{-M-i\alpha}^{M-i\alpha} h(z) dz \xrightarrow{M \rightarrow \infty} \int_{-\infty}^{\infty} \widehat{1_{(0,\infty)}}(-u + i\alpha) \widehat{\mu_X}(u - i\alpha) du \stackrel{(*)}{=} 2\pi [1_{(0,\infty)}(X)].$$

For the integrals (2) and (4) we have:

$\Re e(z) = \pm M$  is fixed and  $-\alpha \leq \Im m(z) \leq 0$ . We also know that for  $0 \leq y \leq \alpha$  the following estimation holds:

$$\begin{aligned} | [e^{i(M-iy)X}] | &\leq | [e^{iMx} |e^{iyX}| ] | = \int_{-\infty}^0 e^{yX} \mu_X(dx) + \int_0^{\infty} e^{yX} \mu_X(dx) \\ &\leq P(X \leq 0) + \int_0^{\infty} e^{\alpha x} \mu_X(dx) = \underbrace{P(X \leq 0)}_{\text{bounded}} + \underbrace{[e^{\alpha x} 1_{\{X>0\}}]}_{\leq [e^{\alpha X}] < \infty} \leq C < \infty. \end{aligned}$$

Hence, on the segment  $\{z \mid \Re e(z) = M, -\alpha \leq \Im m(z) \leq 0\}$  we get

$$|h(z)| \leq \left| \frac{C}{iz} \right| \leq \frac{C}{\sqrt{M^2 + \Im m(z)^2}} \leq \frac{C}{M}$$

and therefore we have for the integral (2):

$$\left| \int_M^{M-i\alpha} h(z) dz \right| = \left| - \int_{M-i\alpha}^M h(z) dz \right| \leq \frac{\alpha C}{M} \xrightarrow{M \rightarrow \infty} 0.$$

Analogously one can show that for the integral (4) we have:

$$\left| \int_{-M-i\alpha}^{-M} h(z) dz \right| \xrightarrow{M \rightarrow \infty} 0.$$

Summarizing our results for the integrals (2) – (4) we get for (\*\*) by taking the limit  $M \rightarrow \infty$ :

$$\int_{-\infty}^{\infty} \widehat{1_{(0,\infty)}}(-u + i\alpha) \widehat{\mu_X}(u - i\alpha) du = \int_{L_{M,\epsilon}} h(z) dz.$$

If we can determine the contribution of the pole in the integral over the subpath  $L_{M,\epsilon}$  when taking the limit  $\epsilon \rightarrow 0$ , we would only have to integrate over the real axis.  $\Rightarrow$  We can determine the expectation in (\*). To solve this problem, we use the coordinate transformation  $z = \epsilon e^{i\theta}$  with  $\theta \in [-\pi, 0]$  (this transformation is quite appropriate because the new coordinates describe a half-circle around the origin with radius  $\epsilon$ .) We then get:

$$\int_{R_\epsilon} h(z) dz = \int_{-\pi}^0 h(\epsilon e^{i\theta}) i \epsilon e^{i\theta} d\theta = \int_{-\pi}^0 \widehat{\mu_X}(\epsilon e^{i\theta}) d\theta,$$

where we used that  $dz = i \epsilon e^{i\theta} d\theta$  and the definition of the function  $h$ . Now taking the limit  $\epsilon \rightarrow 0^+$  we get:

$$\lim_{\epsilon \rightarrow 0^+} \int_{R_\epsilon} h(z) dz = \int_{-\pi}^0 \underbrace{\lim_{\epsilon \rightarrow 0^+} \widehat{\mu_X}(\epsilon e^{i\theta})}_{=\widehat{\mu_X}(0)=1} d\theta = \pi.$$

Using these results we are now able to calculate (\*\*):

$$\begin{aligned} 1 - F_X(0) &= [1_{(0,\infty)}(X)] = \lim_{\substack{m \rightarrow \infty \\ \epsilon \rightarrow 0^+}} \frac{1}{2\pi} \int_{L_{M,\epsilon}} h(z) dz \\ &= \lim_{\substack{m \rightarrow \infty \\ \epsilon \rightarrow 0^+}} \frac{1}{2\pi} \left( \underbrace{\int_{R_\epsilon} h(z) dz}_{\pi} + \underbrace{\int_{\epsilon}^M h(u) + h(-u) du}_{\text{distances } -\epsilon, -M \text{ and } \epsilon, M} \right) \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty h(u) + h(-u) du \quad (*** ) \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} \{ [e^{iuX}] - [e^{-iuX}] \} du \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} \{ \widehat{\mu_X}(u) - \widehat{\mu_X}(-u) \} du \\ &= \left( \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{1}{u} [\sin(uX)] du \right). \end{aligned}$$

$\Rightarrow$  The case  $b = 0$  is shown.

We get the general case by the translation  $X' = X - b$ . If  $X = b$  we have that  $X' = 0$  and we are back in the case  $b = 0$ . Considering

$$[e^{iuX'}] = [e^{iu(X-b)}] = e^{-iub} [e^{iuX}]$$

we get the general case since

$$\begin{aligned} 1 - F_X(b) &= 1 - F_{X'}(0) \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} \{ [e^{iuX'}] - [e^{-iuX'}] \} du \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} \{ e^{-iub} [e^{iuX}] - e^{-iub} [e^{-iuX}] \} du \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} \{ e^{-iub} \widehat{\mu_X}(u) - e^{-iub} \widehat{\mu_X}(-u) \} du. \end{aligned}$$

To complete the proof we still hav to show that there is always an  $\alpha > 0$  with  $[e^{\alpha X}] < \infty$ . Therefore assume that there is not such an  $\alpha$ . Then consider the random variable  $X^a = X \wedge a$ ,  $a > 0$ . For  $X^a$  there is always such an  $\alpha$  since  $X^a$  is bounded. From (\*\*\*) we get:

$$1 - F_{X^a}(0) = \frac{1}{2} + \lim_{M \rightarrow \infty} \frac{1}{\pi} \int_0^M \frac{1}{u} [\sin(uX^a)] du.$$

We have shown the existence of such an  $\alpha$  if

$$\lim_{M \rightarrow \infty} \left[ \int_0^M \frac{1}{u} [\sin(uX)] du - \int_0^M \frac{1}{u} [\sin(uX^a)] du \right] = 0.$$

Of course, this is always fulfilled if  $X < a$ . Using the substitution

$$uX \rightarrow y \quad \Rightarrow \quad du = \frac{1}{X} dy$$

for the left integral and analogously

$$uX^a \rightarrow y \quad \Rightarrow \quad du = \frac{1}{X^a} dy$$

for the right integral as well as interchanging the expectation and the integral we get for the expression in square brackets:

$$\left[ \left( \int_0^{MX} \frac{\sin(y)}{y} dy - \int_0^{Ma} \frac{\sin(y)}{y} dy \right) \cdot 1_{\{X > a\}} \right] = \left[ \int_{Ma}^{MX} \frac{\sin(y)}{y} dy \cdot 1_{\{X > a\}} \right],$$

which tends to 0 as  $M \rightarrow \infty$  by the dominated convergence theorem. This finishes the proof. ■

Summarizing our digression, we can state the following general connection between the characteristic function  $\varphi_X$  of a random variable  $X$  and the corresponding distribution function  $F_X$ :

$$\begin{aligned} P(X \geq b) &= 1 - F_X(b) \\ &= \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} [e^{iub} \widehat{\mu_X}(-u) - e^{-iub} \widehat{\mu_X}(u)] du \\ &= \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} [e^{iub} \varphi_X(-u) - e^{-iub} \varphi_X(u)] du \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty -\frac{1}{iu} [e^{iub} \overline{\varphi_X(u)} - e^{-iub} \varphi_X(u)] du \quad (3.4.25) \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{1}{iu} e^{-iub} \varphi_X(u) + \frac{1}{iu} e^{-iub} \varphi_X(u) du \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[ \frac{e^{-iub} \varphi_X(u)}{iu} \right] du \end{aligned}$$

Now we are finally able to use the obtained result to get a closed-form solution for the Heston model.

(End of Digression 4)



Before the digression we derived the characteristic functions  $\{\varphi_{x(T)}^j(x, \nu, t; u)\}_{j=1,2}$ , where we used the transformation  $x = \ln(S)$  of the asset price. In our particular case the probabilities

$$\{P_j(x \geq \ln[K])\}_{j=1,2} = \{P_j(x, \nu, t; \ln[K])\}_{j=1,2}$$

are of interest. Adjusting formula (3.4.25) to this special case then gives us:

$$P_j(x, \nu, t, \ln[K]) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[ \frac{e^{-iu \ln[K]} \varphi_{x(T)}^j(x, \nu, t; u)}{iu} \right] du. \quad (3.4.26)$$

This finally gives us an up to integration closed-form solution for European calls in the Heston model. We summarize our results in the following theorem:

**Theorem 3.12 (Closed-form Solution for European Calls).**

Let the asset price and its variance evolve according to the following set of stochastic differential equations:

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{\nu(t)}S(t)dW_1(t) \\ d\nu(t) &= \kappa [\theta - \nu(t)] dt + \sigma \sqrt{\nu(t)}dW_2(t) \\ \langle dW_1, dW_2 \rangle &= \rho dt. \end{aligned}$$

Further assume that the market price of volatility risk  $\Phi$  is given by  $\Phi = \lambda \nu(t)$ . Then the arbitrage-free price of a European call written on  $S$  is given by

$$C(t) = C(S(t), \nu(t), t) = S(t) \cdot P_1(S(t), \nu(t), t, \ln[K]) - K \cdot p(t, T) \cdot P_2(S(t), \nu(t), t, \ln[K]),$$

where the probabilities  $P_1$  and  $P_2$  are given by

$$P_j(S(t), \nu(t), t, \ln[K]) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re \left[ \frac{e^{-iu \ln[K]} \varphi_{\ln[S(t)]}^j(\ln[S(t)], \nu(t), t; u)}{iu} \right] du.$$

Thereby the characteristic functions  $\{\varphi_{\ln[S(t)]}^j\}_{j=1,2}$  corresponding to  $\{P_j\}_{j=1,2}$  have the form

$$\varphi_{\ln[S(t)]}^j(\ln[S(t)], \nu(t), t; u) = e^{C_j(T-t; u) + D_j(T-t; u) \cdot \nu(t) + iu \ln[S(t)]}$$

with

$$\begin{aligned} C_j(T-t; u) &= ruiT - t + \frac{a}{\sigma^2} \left\{ (b_j - \rho\sigma ui + d_j)(T-t) - 2 \ln \left[ \frac{1 - g_j e^{d_j(T-t)}}{1 - g_j} \right] \right\}, \\ D_j(T-t; u) &= \frac{b_j - \rho\sigma ui + d_j}{\sigma^2} \cdot \left[ \frac{1 - e^{d_j(T-t)}}{1 - g_j e^{d_j(T-t)}} \right], \end{aligned}$$

where

$$\begin{aligned} g_j &= \frac{b_j - \rho\sigma ui + d_j}{b_j - \rho\sigma ui - d_j}, \\ d_j &= \sqrt{(\rho\sigma ui - b_j)^2 - \sigma^2(2u_j ui - u^2)} \end{aligned}$$

and

$$u_1 = \frac{1}{2}, \quad u_2 = -\frac{1}{2}, \quad a = \kappa\theta, \quad b_1 = \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda.$$

### 3.5 Calibration and numerical Results

Having achieved a closed-form solution for European calls, we are now able to calibrate the Heston model to given market prices of traded plain vanilla calls. We follow [Kruse & Nögel, 2005] and [Mikhailov & Nögel, 2003]. In digression 2 we already described how a calibration procedure is done in general. From theorem 3.12 we can gather that the unknown vector of parameters

$$\Theta = (\kappa, \theta, \sigma, \rho, \nu(0), \lambda) \in (0, \infty] \times [0, \infty] \times (0, \infty] \times [-1, 1] \times (0, \infty) \times \quad =: \Xi$$

is needed for the calculation of the closed-form solution for a European call at time  $t = 0$ . To be able to adequately formulate the calibration task we give at first some notation:

- Let  $(C_M(K_i, T_j))_{i=1, \dots, N; j=1, \dots, M}$  be a set of market prices of European calls with strikes  $(K_i)_{i=1, \dots, N}$  and maturities  $(T_j)_{j=1, \dots, M}$  at time  $t = 0$ .
- Let  $C_H(0, S(0), K_i, T_j, \Theta)$  denote the theoretical Heston price of a European call with strike  $K_i$  and  $T_j$  at time  $t = 0$  when the vector of parameters  $\Theta$  is used.

Then a first suggestion for the calibration task is to solve the least-squared error problem

$$\min_{\Theta \in \Xi} \sum_{i=1}^N \sum_{j=1}^M \{C_M(K_i, T_j) - C_H(0, S(0), K_i, T_j, \Theta)\}^2 \quad (3.5.1)$$

w.r.t. the nonlinear constraint  $2\kappa\theta \geq \sigma^2$ .

One can improve this calibration procedure by adding weight factors  $(\omega_{i,j})_{i=1, \dots, N; j=1, \dots, M}$  and make it more stable by adding a penalty term which is for example given by the distance between of the actual parameter vector  $\Theta$  and an initial parameter vector  $\Theta_0$ . This is obviously a good choice if one wants to calibrate the model daily to market data (just use the results from the day before as  $\Theta_0$ ). Further, it turns out that the choice of the weight factors is crucial for good calibration results (see [Mikhailov & Nögel, 2003]). The corresponding least-squared error problem then reads:

$$\min_{\Theta \in \Xi} \left( \sum_{i=1}^N \sum_{j=1}^M \omega_{i,j} \{C_M(K_i, T_j) - C_H(0, S(0), K_i, T_j, \Theta)\}^2 + \|\Theta - \Theta_0\| \right) \quad \text{w.r.t. } 2\kappa\theta \geq \sigma^2. \quad (3.5.2)$$

#### Remarks.

- Calibrating this way also deals with the incompleteness of the market in the Heston model. Since the parameter  $\lambda$  (which determines the market price of risk due to  $\Phi = \lambda\nu$ ) is part of the parameter vector  $\Theta$ , we can choose the martingale measure which is closest to the market after the calibration.
- We can also go through the calibration procedure using the risk-neutral vector of parameters

$$\Theta^* = (\kappa^*, \theta^*, \sigma, \rho, \nu(0)) \in (0, \infty] \times [0, \infty] \times (0, \infty] \times [-1, 1] \times (0, \infty) =: \Xi^*$$

with  $\kappa^* = \kappa + \lambda$  and  $\theta^* = \frac{\kappa\theta}{\kappa^*}$ . The least-squared error problems then stay the same, just replace  $\Theta$  by  $\Theta^*$ ,  $\Xi$  by  $\Xi^*$  and  $\Theta_0$  by  $\Theta_0^*$ .

As next we explain how the optimization is done in detail. The calibration task (3.5.2) is clearly a nonlinear programming problem with a nonlinear constraint. Since the objective function is not convex and since usually there exist many local extrema, the minimization problem is not easy at all. [Mikhailov & Nögel, 2003] therefore suggested to use both local and global optimizers:

- **Local (deterministic) algorithms:**

To apply these types of algorithms one has to choose a good initial guess for the parameter vector  $\Theta_0 \in \Xi$ . The algorithm then determines the optimal direction and the stepsize and is moving downhill to the minimum of the objective function. There are a lot of algorithms available for these types of optimization problems; usually they are based on simplex or some kind of gradient method (steepest descent). But most of these methods always stop if the difference between a new point and a previous point is smaller than a fixed value. Therefore there is always the risk to end up in a local minimum. As a consequence a good initial guess is crucial. → Such an optimizer is predestined for the daily recalibration of a model when the volatility surface has not changed so much.

- **Stochastic algorithms:**

In contrast to the local optimizers the initial guess is irrelevant in the concept of stochastic optimization. Therefore stochastic optimizers should be used for the first calibration of a model when the initial guess is probably not very close to the optimal value. We consider as an example the simulated annealing algorithm. This algorithm chooses the direction and stepsize randomly → it searches globally and not only locally. It moves always downhill but may accept an uphill move with a certain probability  $P_T$  which depends on the annealing parameter  $T$ . During the optimization process the annealing parameter  $T$  is gradually reduced (annealing process). There are convergence theorems that ensure that such an algorithm always ends up in the global minimum if the annealing process is slow enough. A disadvantage of these stochastic algorithms is that their implementation is more time consuming than that of local algorithms.

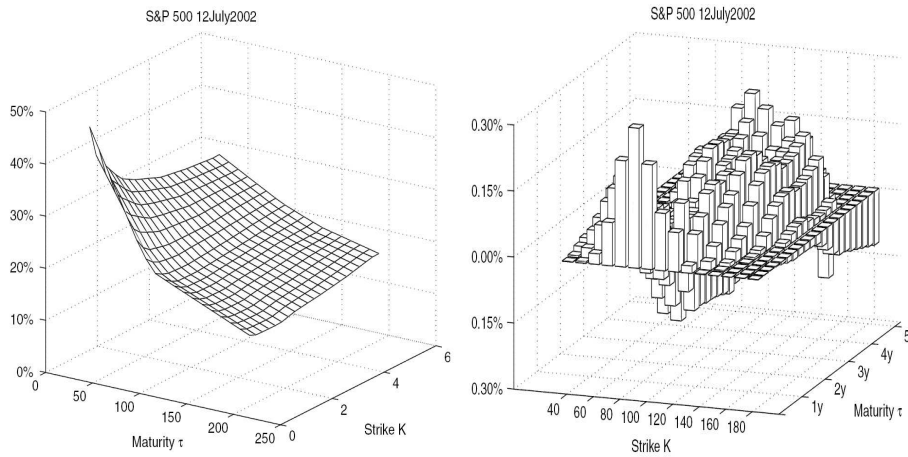


Figure 11: Volatility Surface of the S&P 500 Index and the relative Errors after Calibration to Heston's Model

Numerical results underline the quality of Heston's model. Figure 11 shows the volatility surface of the S&P 500 Index on 12 July, 2002 and the relative errors after the calibration of the Heston model (these are the difference of the volatility surface obtained from the market prices and the volatility surface obtained from the Heston prices calculated with the calibrated parameter vector). One can see that the volatility surface is fitted with high precision. The maximum error is less than 0.15% for at-the-money calls. A slightly larger relative error can be observed for options which are far out-the-money or far in-the-money.

### 3.6 The Limits of Heston

The Heston model has a far field of applications. It can easily be extended to dividend-paying assets and stochastic interest rates (see [Heston, 1993]) and is among other things applied to the following options and option pricing tasks by implementing suitable variants of it (see [Korn, 2006]):

- standard European options
  - pricing (via closed-form, Monte-Carlo simulation and finite differences)
  - calculation of Greeks (via closed-form and finite differences)
  - calibration using local and global optimizers
- forward-starting options
  - pricing (via closed-form, Monte-Carlo simulation and finite differences)
  - calibration to traded options using local and global optimizers
- structured options like cliquets
  - pricing (via Monte-Carlo simulation and finite differences)
  - calculation of Greeks (via Monte-Carlo simulation and finite differences).

The Heston model was thereby assessed according to how it fits given market option prices across strikes and maturities, i.e. how it fits the volatility surface. However, the pricing of more recent exotic options such as *reverse cliquets*, *Napoleons* or *accumulators* is more dependent on the assumptions made for the future dynamics of implied volatilities than on today's vanilla option prices. So for these types of options also the Heston model reaches its limits and gives wrong option prices. In this section we give evidence for that by highlighting some structural features of the dynamic properties of the Heston model. We follow [Bergomi, 2004] and [Korn, 2006].

#### 3.6.1 Example: Napoleon and its Dependence on Future Smiles

In the Black-Scholes model implied volatilities for different strikes are equal and fixed. Over the years several alternative models have been developed in order to fit the volatility surface. The first models were local volatility models which we explained in chapter 2. The formula of Dupire thereby shows how today's market prices are fitted. As next generation stochastic volatility models like Heston's model emerged. In chapter 3.5 we showed how excellent the calibration to the Black-Scholes implied volatilities is.

This capability is a desirable feature of any smile model. The model price then incorporates by construction the cost of trading vanilla options to hedge the exotic option's vega risk - at least for the initial trade. Otherwise the price has to be manually adjusted to reflect hedging costs;

that is why the difference between market and model prices of vanilla options is used for the hedge. This may be sufficient if the vega hedge is stable (which is usually the case for barrier options).

Most of the recent exotic options such as Napoleons and reverse cliquets (see the review article of [Jeffery, 2004]) require rebalancing of the vega hedge when the underlying or its implied volatilities move substantially. To ensure that future hedging costs are priced-in correctly, the model has to be designed so that it incorporates from the start a dynamic for the implied volatilities that is consistent with the historically experienced one.

Stated differently, for this type of options the gammas

$$\Gamma_{\hat{\sigma}\hat{\sigma}} = \frac{\partial^2 C}{\partial \hat{\sigma}^2} \quad \text{and} \quad \Gamma_{S\hat{\sigma}} = \frac{\partial^2 C}{\partial S \partial \hat{\sigma}}$$

are sizeable and a suitable model needs to price in a theta to match these gammas. In the view of Lorenzo Bergomi (see the article [Bergomi, 2004]), this issue is more important than the model's ability to reproduce today's volatility surface.

As an illustration let us consider the example of a Napoleon option with a maturity of 6 years. This option has the following (payoff) structure:

- Initial invest: Fixed amount  $N_0$ .
- End of years 1 and 2: Guaranteed coupon of  $C = 6\%$ .
- End of years 3,4 and 5: Payoff =  $\max\left\{8\% + \min_{i=1}^{12} r_i; 0\right\}$ , where  $\{r_i\}_{i=1,\dots,12}$  are the monthly performances of the Eurostoxx 50 index observed for each year.
- End of year 6: The initial invest  $N_0$  is paid back.

The payoffs for the last 4 years is thereby designed so that their value is relatively small in order to finance the large fixed initial coupons which we remove from the option in what follows.

Figure 12 shows on the left-hand side the Black-Scholes value of the Napoleon at time  $t = 0$  as a function of volatility. As we can see, the Napoleon is in essence a put option on long (one-year) forward volatility (the graph follows the value of a regular put option), for which no time value has been appropriated for in the Black-Scholes price. This is due to the fact that when the volatility rises the vega rises too and this means that we have to buy volatility then, i.e.

$$\Gamma_{\hat{\sigma}\hat{\sigma}} = \frac{\partial^2 C}{\partial \hat{\sigma}^2} > 0$$

and therefore we need a theta matching  $\Gamma_{\hat{\sigma}\hat{\sigma}}$  which is not given in the Black-Scholes model.

Now let us move to the end of the first month of year three. The right-hand side of Figure 12 shows the vega of the coupon of year three as a function of the spot price under the assumption that  $S(0) = 100$ . It is a decreasing function of the spot price and goes to zero for low spot values because then the coupon becomes worthless. So as the spot decreases the option seller will need to buy back vega.

However, moves in spot prices are historically negatively correlated with moves in implied volatilities. This results then in a negative profit and loss to the seller which is not factored in by the Black-Scholes price, i.e. there is no theta matching

$$\Gamma_{S\hat{\sigma}} = \frac{\partial^2 C}{\partial S \partial \hat{\sigma}} < 0.$$

So our example of the Napoleon option shows an equity price model should have the desirable property to pay a theta to offset the gammas  $\Gamma_{\hat{\sigma}\hat{\sigma}}$  and  $\Gamma_{S\hat{\sigma}}$ . Additionally a model should incorporate the one-month forward skew contribution which is needed due to the fact that we need to determine the worst monthly performance of the Eurostoxx 50 index over three years for the value of the coupons. The future performance of the Eurostoxx 50 index seen from the yearly coupon dates obviously depends strongly on the forward volatility surface. So it would be desirable to be able to independently calibrate today's market smile and specify its future dynamics.

In the following sections we will show that the Heston model is not able to guarantee these objectives at the same time and that therefore a more advanced equity price model is needed.

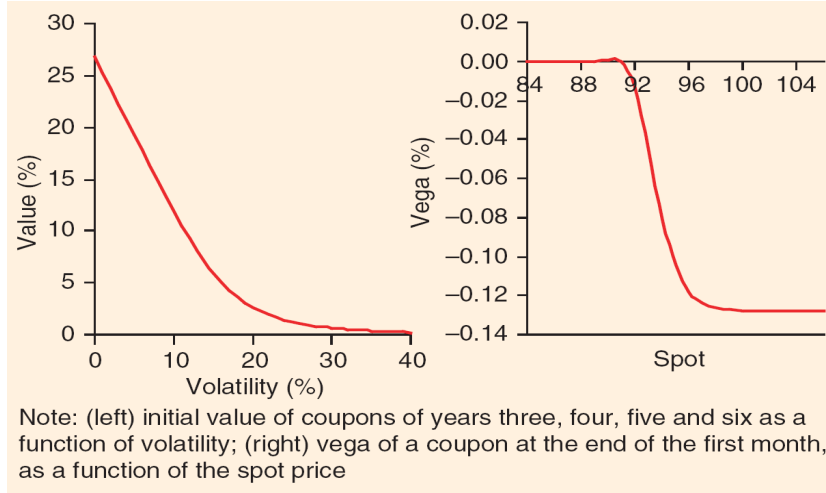


Figure 12: Price and Vega of Napoleon

### 3.6.2 Static Properties of the Heston Model

Starting with the risk-neutral pricing PDE (3.3.10) the Heston model has the five unknown parameters  $\nu, \theta^*, \rho, \sigma$  and  $\kappa^*$ . In the following we skip the star and write  $\theta$  and  $\kappa$ . Among these parameters,  $\kappa$  plays a special role:

$$\tau := \frac{1}{\kappa} \quad \text{is a cutoff that separates short and long maturities.}$$

In the Heston model, implied volatilities are a function of  $\nu$  and the strike price:

$$\hat{\sigma} = \sigma_{\text{imp}} = f\left(\frac{K}{F}, \nu\right),$$

where  $F = F_{t,T} = S(t)e^{r(T-t)}$  is the forward price of the stock at time  $t$  in the risk-neutral world assuming constant interest rates.

In [Bergomi, 2004] it is shown by a perturbation of the risk-neutral equation (3.3.10) at first

order that the following expressions for the skew  $\frac{d\hat{\sigma}}{d \ln K}$  and the at-the-money-forward volatility<sup>8</sup> (ATMF volatility)  $\hat{\sigma}_F$  are valid:

- $T \ll \tau$  : At order zero in  $T$ :

$$\hat{\sigma}_F = \sqrt{\nu} \quad (3.6.1)$$

$$\left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\rho\sigma}{4\sqrt{\nu}} \quad (3.6.2)$$

- $T \gg \tau$  : At order zero in  $\frac{1}{T}$ :

$$\hat{\sigma}_F = \sqrt{\theta} \left( 1 + \frac{\rho\sigma}{4\kappa} \right) + \frac{\sqrt{\theta}}{2\kappa T} \left( \frac{\nu - \theta}{\theta} + \frac{\rho\sigma}{4\kappa} \frac{\nu - 3\theta}{\theta} \right) \quad (3.6.3)$$

$$\left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\rho\sigma}{2\kappa T \sqrt{\theta}} \quad (3.6.4)$$

From equations (3.6.1), (3.6.2), (3.6.1) and (3.6.2) we can conclude the following:

- Since  $\sigma, \nu, \kappa, T$  and  $\sigma$  are bigger than zero, the parameter  $\rho$  is needed to produce the negative skew.
- Close to maturity, i.e. for  $T \ll \tau$ ,  $\sqrt{\nu}$  equals the ATMF volatility.
- Far away from maturity, i.e. for  $T \gg \tau$ ,  $\sqrt{\theta}$  determines the ATMF volatility.
- The short-term skew is proportional to  $\frac{1}{\sqrt{\nu}}$ .
- The long-term skew is independent of  $\nu$ , proportional to the inverse squareroot of the long term mean  $\frac{1}{\sqrt{\theta}}$  and decreases like  $\frac{1}{T}$ .

In the following sections we will use the relations (3.6.1), (3.6.2), (3.6.1) and (3.6.2) and the conclusions above to show some limitations of the Heston model.

As next we deduce a further static property a the Heston model from variance swaps. Consider the volatility  $\hat{\sigma}_{VS}$  corresponding to a variance swap. To be able to deal with this notion we first give the definition of a variance swap.

**Definition 3.13 (Variance Swap):**

A variance swap is a forward contract on the annualized variance. Its payoff at expiration is equal to

$$N \cdot (\sigma_R^2(S) - K_{\text{Var}}) ,$$

where  $\sigma_R^2(S)$  is the realized stock variance (quoted in annual terms) over the life time of the contract,

$$\sigma_R^2(S) := \frac{1}{T} \int_0^T \sigma^2(t) dt ,$$

---

<sup>8</sup>The notion of at-the-money forward volatility just means the usual implied (Black-Scholes) volatility for a strike which is equal to the forward price for the maturity consireded, i.e.

$$K = F(t, T) = S(t)e^{r(T-t)} .$$

$K_{\text{Var}}$  is the delivery price for variance and  $N$  is the notional amount of the swap. So the holder of the variance swap at expiration receives  $N$  units for every point by which the stock's realized variance has exceeded the variance delivery price.

**Remark.** For practical purposes one cannot deal with the continuously realised variance

$$\text{Var}(S) := \sigma_R^2(S) := \frac{1}{T} \int_0^T \sigma^2(t) dt,$$

since we only have daily data and the volatility itself is no traded asset. To solve this problem we use a notion very similar to the  $N$ -day historic volatility (compare definition 3.1). Therefore let  $\{r(n)\}_{n=1, \dots, M}$  denote the time series of the log returns  $r(n) := S(n)/S(n-1)$  of an asset or an index. Then the **discrete realised variance** at the discrete time point  $N$  is for  $M \geq N$  is defined as the sample variance of the log returns scaled by a factor of  $N/T$ :

$$\text{Var}_N(S) := \frac{N}{(N-1)T} \sum_{n=1}^N (r(n) - \bar{r}_N)^2 \quad \text{with} \quad \bar{r}_N = \frac{1}{N} \sum_{n=1}^N r(n).$$

Thereby the scaling factor  $N/T$  ensures that these quantities are annualized if the maturity  $T$  is expressed in years respectively daily if the maturity  $T$  is expressed in days.  $\text{Var}_N(S)$  is an unbiased variance estimation for  $\sigma(t)$ ; in [Brockhaus & Long, 2000] it is shown that  $\lim_{N \rightarrow \infty} \text{Var}_N(S) = \text{Var}(S)$ .

A variance swap can be valued in exactly the same way as any other derivative security. Its price is the expected present value of the future payoff in the risk-neutral world:

$$C_{\text{VS}}(t) = \left[ e^{-r(T-t)} N (\sigma_R^2(S) - K_{\text{Var}}) \right] = e^{-r(T-t)} N (\sigma_R^2(S) - K_{\text{Var}}).$$

Thus for calculating the value of a variance swap we only need to know

$$\hat{\sigma}_{\text{VS}}^2 := [\sigma_R^2(S)]$$

which we call the mean value of the underlying variance resp. the variance swap variance.  $\hat{\sigma}_{\text{VS}}$  is then the variance swap volatility.

We get for the value of  $\hat{\sigma}_{\text{VS}}^2$  when we assume the volatility dynamics of the Heston model:

$$\hat{\sigma}_{\text{VS}}^2 = [\sigma_R^2(S)] = \frac{1}{T} \int_0^T [\sigma^2(t)] dt = \frac{1}{T} \int_0^T [\nu(t)] dt.$$

In [Zhang, 2004] it is shown that the expectation of a Cox-Ingersoll-Ross process equals

$$[\nu(t)] = \theta + (\nu(0) - \theta)e^{-\kappa t}.$$

So we get:

$$\begin{aligned} \hat{\sigma}_{\text{VS}}^2 &= \frac{1}{T} \int_0^T \theta + (\nu(0) - \theta)e^{-\kappa t} dt \\ &= \frac{1}{T} \left[ \theta t - \frac{(\nu(0) - \theta)}{\kappa} e^{-\kappa t} \right]_0^T \\ &= \frac{1}{T} \left\{ \theta T - \frac{(\nu(0) - \theta)}{\kappa} e^{-\kappa T} + \frac{(\nu(0) - \theta)}{\kappa} \right\} \\ &= \theta + (\nu(0) - \theta) \frac{1 - e^{-\kappa T}}{\kappa T}. \end{aligned} \tag{3.6.5}$$



This means that in the Heston model the dynamics are seen for short maturities (remember  $\tau = \frac{1}{\kappa}$ ).

**Remark.** Variance swap volatilities can be used as a market measure for the quality of implied volatilities since they are frequently traded and so they allow to draw conclusions of the realistic behaviour of volatilities. Figure 13 shows the ATM implied volatility of the Eurostoxx 50 index from 03/12/1999 to 03/12/2004 for options with maturity of one year compared to the variance swap volatility  $\hat{\sigma}_{VS}$  determined by equation (3.6.5) with a maturity of one year, where the parameters  $\kappa$ ,  $\theta$  and  $\nu(0)$  are taken from a daily calibration of the Heston model. One can see that the calibration is satisfactory.

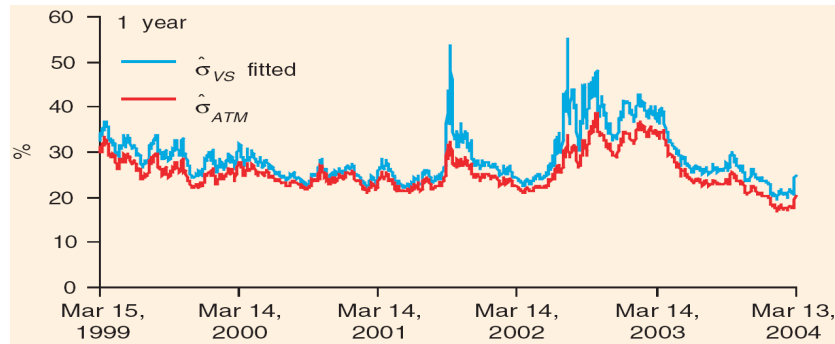


Figure 13: Variance Swap Volatility and 1-Year ATM Volatility

### 3.6.3 Dynamics of the Heston Parameters

In the Heston model the asset price  $S$  and the variance  $\nu$  are dynamic and the parameters mean reversion level  $\theta$ , mean reversion speed  $\kappa$ , volatility of volatility  $\sigma$  and correlation  $\rho$  are supposed to be constant. As already explained these parameters are needed for option pricing and are determined by calibrating the model to market prices of traded options.

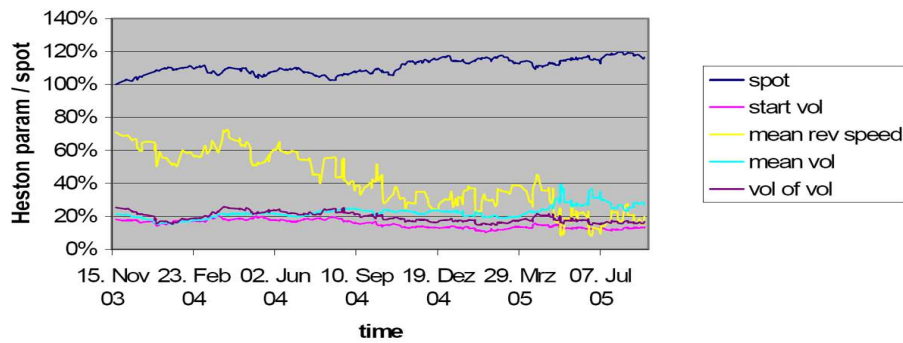


Figure 14: Time Evolution of Heston Parameters for SPX Index

In figure 14 we have plotted the parameters  $\nu(0)$ ,  $\kappa$ ,  $\theta$  and  $\sigma$  obtained from calibration of the

SPX Index to the Heston model from November 2003 to July 2005. The spot price is also plotted for reference. This figure shows that  $\kappa$ ,  $\theta$  and  $\sigma$  vary over time and are empirically not constant which is a contradiction to the assumptions made by the Heston model.

In figure 15 we have plotted on the left-hand side the graphs of the stock price and the volatility as a quantitative measure of the correlation  $\rho$  and on the right-hand side the correlation  $\rho$  between stock price and volatility obtained from calibration of the Eurostoxx 50 Index from March 1999 to March 2004 to the Heston model. This figure shows that the stock-volatility correlation is very high and that  $\rho$  is not constant, too, but fairly stable and does not seem to be correlated with the other parameters. So we can conclude that the fact that the parameters  $\kappa$ ,  $\theta$ ,  $\sigma$  and  $\rho$  are dynamic is not priced-in by the Heston model!

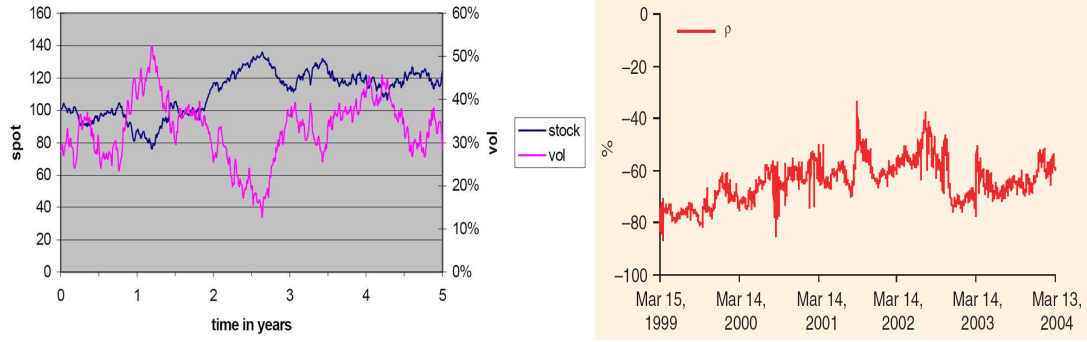


Figure 15: Quantitative Stock-Volatility Correlation and Time Evolution of  $\rho$

Among all these parameters,  $\sigma$  is the most interesting one. Figure 16 shows the values of the  $\nu(0)$  and  $\sigma$  again obtained from calibration of the Heston model to the Eurostoxx 50 Index from March 1999 to March 2004. Thereby the graph of  $\nu(0)$  is superimposed with a scale 10 times larger. We see that  $\sigma$  varies substantially and seems to be closely correlated with  $\nu(0)$ . This shows a further mis-specification of the Heston model. We showed that the short-term skew in the Heston model is approximately

$$\left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\rho\sigma}{4\sqrt{\nu}} = \frac{\rho\sigma}{4\hat{\sigma}_F},$$

i.e. inversely proportional to the at-the-money volatility. Our empirical observations have shown that

$$\sigma \propto \nu.$$

So we have for the empirical market skews:

$$\left. \frac{d\hat{\sigma}_{\text{Market}}}{d \ln K} \right|_F = \frac{\rho\nu}{4\sqrt{\nu}} = \frac{\rho\sqrt{\nu}}{4} = \frac{\rho\hat{\sigma}_F}{4},$$

i.e. the empirical observations suggest that market skews are proportional to at-the-money volatilities.

So in this respect the Heston model is mis-specified, since it is not pricing in the observed correlation between  $\nu$  and  $\sigma$ . This correlation is very visible in the graphs, mostly for extreme events. However this fact is even visible for not so extreme events. For example, the daily

variations of  $\nu$  and  $\sigma$  measured from March 15, 1999 to September 10, 2001 have a correlation of 59%.

So the main result of this historical analysis is that the volatility of volatility  $\sigma$  and the variance  $\nu$  are closely related.

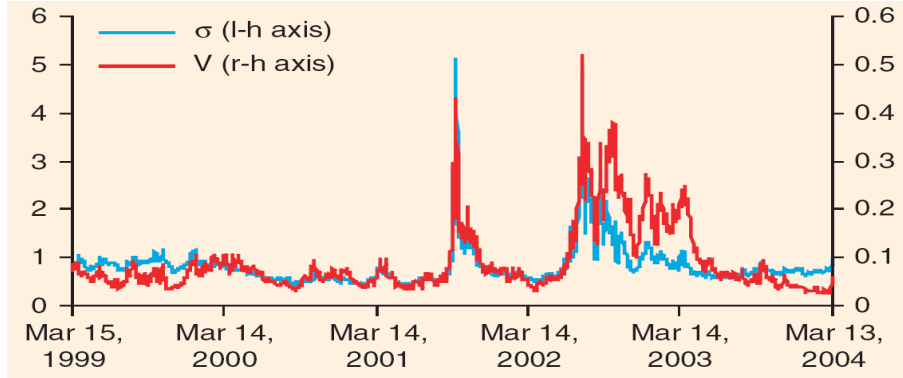


Figure 16: Time Evolution of Vol-of-Vol compared to Start Vol

### 3.6.4 Dynamics of implied Volatilities

In this section we consider the implied volatilities generated by the Heston model and compare them with the historical dynamics. In the Heston model, the implied volatility dynamics are determined by construction by that of  $S$  and  $\nu$ . We use daily values of the couple  $(S, \nu)$  and investigate whether their dynamics are consistent with the model specification. The empirical results that are stated in what follows, are taken from [Bergomi, 2004].

To examine the historical dynamics, we discuss the following averages for the Eurostoxx 50 index from 03/12/1999 to 03/12/2004, which in theory should all be equal to one:

$$R_S = \left\langle \frac{\delta S^2}{S^2 \nu \delta t} \right\rangle = 0.75 \quad R_\nu = \left\langle \frac{\delta \nu^2}{\sigma^2 \nu \delta t} \right\rangle = 0.4 \quad R_{S\nu} = \left\langle \frac{\delta S \delta \nu}{\rho \sigma S \nu \delta t} \right\rangle = 0.6.$$

Thereby brackets denote historical averages using daily variations.

With these numbers we are now able to estimate the ratios

$$\frac{\sigma_{\text{realised}}}{\sigma_{\text{implied}}} = \sqrt{R_\nu} = 0.63 \quad \text{and} \quad \frac{\rho_{\text{realised}}}{\rho_{\text{implied}}} = \frac{R_{S\nu}}{\sqrt{R_\nu R_S}} = 1.1.$$

This means that the calibration on market smiles of the Heston model overestimates the volatility of volatility  $\sigma$  by 40% and that the value of the spot/volatility correlation is estimated with acceptable accuracy.

In contradiction to what we would expect,  $R_S$  is remarkably different from one. This means that short implied volatilities have overestimated the historical volatility by 13% in our example here.

But it is possible that these long-time averages are immoderately affected by extreme events. So we now look at running monthly averages. Figure 17 shows the result of the six following

quantities:

$$\begin{aligned}\nu_{\frac{\delta S}{S}}^{\text{real}} &= \left\langle \frac{\delta S^2}{S^2} \right\rangle & \text{and} & \quad \nu_{\frac{\delta S}{S}}^{\text{impl}} = \langle \nu \delta t \rangle \\ \nu_{\delta \nu}^{\text{real}} &= \langle \delta \nu^2 \rangle & \text{and} & \quad \nu_{\delta \nu}^{\text{impl}} = \langle \sigma^2 \nu \delta t \rangle \\ C_{\frac{\delta S}{S} \delta \nu}^{\text{real}} &= \left\langle \frac{\delta S}{S} \delta \nu \right\rangle & \text{and} & \quad C_{\frac{\delta S}{S} \delta \nu}^{\text{impl}} = \langle \rho \sigma \nu \delta t \rangle .\end{aligned}$$

Now the brackets denote running monthly averages.

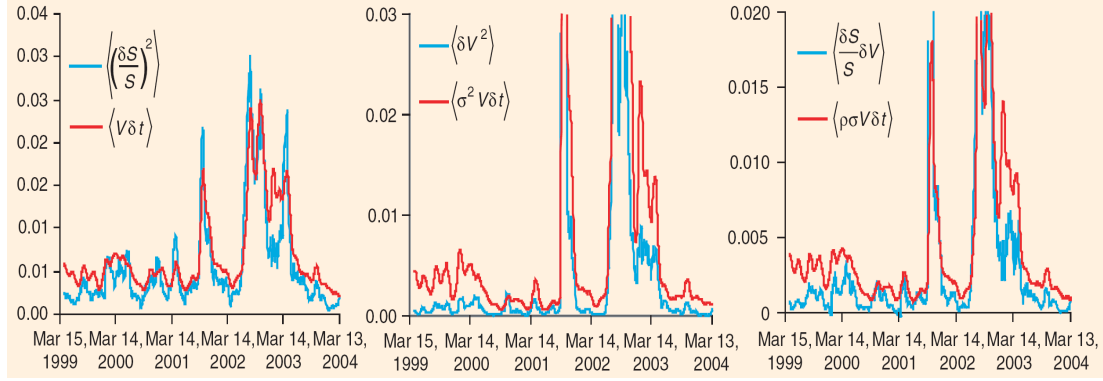


Figure 17: Running monthly Averages of  $\nu_{\frac{\delta S}{S}}^{\text{real}}$  and  $\nu_{\frac{\delta S}{S}}^{\text{impl}}$ ,  $\nu_{\delta \nu}^{\text{real}}$  and  $\nu_{\delta \nu}^{\text{impl}}$ ,  $C_{\frac{\delta S}{S} \delta \nu}^{\text{real}}$  and  $C_{\frac{\delta S}{S} \delta \nu}^{\text{impl}}$

From the figure we see that even during normal market conditions the difference between the realised and the implied quantities is substantial. For example, we get the following numbers using monthly averages estimated on data from 03/15/1999 to 09/10/2001:

$$R_S = 0.73 \quad , \quad R_\nu = 0.30 \quad , \quad R_{S\nu} = 0.44 .$$

This corresponds to the following ratios:

$$\frac{\sigma_{\text{realised}}}{\sigma_{\text{implied}}} = 0.54 \quad \text{and} \quad \frac{\rho_{\text{realised}}}{\rho_{\text{implied}}} = 0.95 ,$$

which shows again that  $\rho$  is well-captured by calibrating the Heston model to market smiles and that  $\sigma$  is overestimated by roughly a factor 2.

This means that the Heston model is pricing in a volatility of volatility for one-month at-the-money volatilities that is twice as large as the historical value. Therefore future vega re-hedging costs are not properly priced-in.

It also implies that in the Heston model hedges are inefficient for short-term maturities. The main result in this section of the historical analysis is that the value of the volatility of volatility  $\sigma$  determined from calibration on market smiles is larger by a factor of two than its historical value. This has structural reasons:

We have only one device in the Heston model - namely the volatility of volatility  $\sigma$  - to achieve two different objectives, the one static and the other dynamic:

- Fit the short-term skew.

- Drive the dynamics of implied volatilities in a way that is consistent with their historical behaviour.

It is natural that one device is unable to fulfill both objectives at the same time. This is a structural limitation of any one-factor stochastic volatility model and here in particular of the Heston model. This shows the need for a more advanced equity price model that can guarantee both objectives.

### 3.6.5 Forward-starting Options

In this section we describe the performance of the Heston model for forward-starting options. The payoff of this type of options depends on a single future date and so the examination of the Heston prices for these options gives a flavour of how the Heston model works for options that strongly depend on future smiles and not only on one future date.

Consider a relative one-period forward call option with maturity  $T + \eta$ , starting point  $T$ , strike  $K$  and payoff

$$P_{\text{RFWS}}(S(T + \eta), S(T)) = \max \left\{ \frac{S(T + \eta)}{S(T)} - K; 0 \right\}. \quad (3.6.6)$$

From time  $T$  on, these options can be priced by the Black-Scholes model since at this point the payoff is known due to the fact that  $S(T)$  is then a constant. Set  $X(T + \eta) := \frac{S(T + \eta)}{S(T)}$ , then the Black-Scholes formula (compare theorem 1.2) applied to

$$P_{\text{RFWS}}(X(T + \eta)) = \max \{X(T + \eta) - K; 0\}$$

yields keeping in mind that  $X(T) = 1$ :

$$C_{\text{RFWS}}^{\text{BS}}(T) = C_{\text{RFWS}}^{\text{BS}}(T, S(T), \sigma_{\text{BS}}, K) = \Phi(d_1(T)) - K \cdot \Phi(d_2(T)) \cdot e^{-r\eta}$$

with

$$d_1(T) = \frac{\ln\left(\frac{1}{K}\right)(r + 0.5\sigma_{\text{BS}}^2)\eta}{\sigma_{\text{BS}}\sqrt{\eta}} \quad \text{and} \quad d_2(T) = d_1(T) - \sigma_{\text{BS}}\sqrt{\eta}.$$

The time- $t$  price is then given by the expectation of the discounted value of the option at time  $T$ , i.e.

$$\begin{aligned} C_{\text{RFWS}}^{\text{BS}}(t, S(t), \sigma_{\text{BS}}, K) &= \left[ e^{-r(T-t)} \cdot C_{\text{RFWS}}^{\text{BS}}(T, S(T), \sigma_{\text{BS}}, K) \right] \\ &= \left[ e^{-r(T-t)} \left( \Phi(d_1(T)) - K \cdot \Phi(d_2(T)) \cdot e^{-r\eta} \right) \right] \\ &= e^{-r(T-t)} \cdot \left( \Phi(d_1(T)) - K \cdot e^{-r\eta} \cdot \Phi(d_2(T)) \right). \end{aligned} \quad (3.6.7)$$

In [Kruse & Nögel, 2005] it is shown that the time- $t$  price of a relative forward starting call option with payoff (3.6.6) in Heston's model is given by

$$C_{\text{RFWS}}^H(t, \nu(t), S(t), K) = e^{-r(T-t)} \left( \hat{P}_1(t, \nu(t)) - K e^{-r\eta} \hat{P}_2(t, \nu(t)) \right) \quad (3.6.8)$$

with

$$\hat{P}_j(t, \nu(t)) := \int_0^\infty P_j(1, v, T, \ln[K]) \cdot f(v|\nu(t)) dv$$

where the  $P_j$ ,  $j = 1, 2$  are equal to the Heston probabilities in theorem 3.12 and

$$f(v|\nu(t)) = \frac{B}{2} e^{-(Bv+\Lambda)/2} \left( \frac{Bv}{\Lambda} \right)^{(R/2-1)/2} I_{R/2-1}(\sqrt{\Lambda Bv})$$

with

$$\Lambda = B e^{-\kappa(T-t)} \nu(t) \quad , \quad B = \frac{4\kappa}{\sigma^2} \left( 1 - e^{-\kappa(T-t)} \right)^{-1} \quad \text{and} \quad R = \frac{4\kappa\theta}{\sigma^2} ,$$

while  $I_{R/2-1}(\cdot)$  is a modified Bessel function of the first kind.

Now we are able to apply the concept of implied volatility to the price of a forward starting call in Heston's model. We fix maturity  $T + \theta$  and require that  $C_{\text{RFWS}}^H$  - which was computed using (3.6.8) - is equal to the Black Scholes price  $C_{\text{RFWS}}^{BS}$  given by equation (3.6.7) of the option for a given set of strike prices  $(K_i)_{i=1, \dots, N}$ , i.e.

$$C_{\text{RFWS}}^H(t, \nu(t), S(t), K_i) \stackrel{!}{=} C_{\text{RFWS}}^{BS}(t, S(t), \sigma_{\text{BS}}^{\text{imp}}, K_i) \quad \text{for } i = 1, \dots, N .$$

Then  $\hat{\sigma}(K_i) := \sigma_{\text{BS}}^{\text{imp}}(K_i)$  is the value of the Black-Scholes volatility for which the Black-Scholes option value matches the Heston option value.

The plot of  $\hat{\sigma}(K)$  as a function of the strike price is then called

- **forward smile** (resp. forward skew) for  $T > 0$  and
- **today's smile** (resp. today's skew) for  $T = 0$ .

Figure 18 shows the forward smile for  $T = 0.25$  (in three months),  $T = 0.5$  (in six months) and  $T = 1$  (in one year) calculated using the typical values  $\nu(0) = 0.1$ ,  $\theta = 0.1$ ,  $\sigma = 1$ ,  $\rho = -0.7$  and  $\kappa = 2$  for the parameters and the values  $\eta = 0.25$  (three months) and  $\eta = 1$  (one year) for the distance between starting point and maturity. Today's smile is also plotted for reference.

Note that forward smiles are more convex than today's smile. This is due to the fact that the price of a forward call (compare formula (3.6.7)) is an increasing function of its implied Black-Scholes volatility; uncertainty of future implied volatility then increases the option price.

Further, one can obtain from figure 18 that the forward smiles collapse on to a single curve for  $T \gg 6$  months in our example. This is because when  $T$  is more distant, the distribution of the variance  $\nu$  in the Heston model becomes stationary (see also the explanation below).

The graphs also show that the increased convexity of forward smiles w.r.t. today's smile is larger for strikes  $K > 100\%$  than for strikes  $K < 100\%$ . This can be explained by the dependence of the skew on the level of at-the-money-volatility. Therefore remember the properties of the skew in the Heston model:

$$\left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\rho\sigma}{4\sqrt{\nu}} = \frac{\rho\sigma}{4\hat{\sigma}_F} \quad \text{for } T \ll \frac{1}{\kappa} ,$$

i.e. the short-term skew is inversely proportional to the level of at-the-money volatility. This dependence shows that implied volatilities for  $K > 100\%$  will move more than those for strikes  $K < 100\%$ .

While the forward smile is a global measure of the distribution of implied volatilities at a forward starting date, it is instructive to look at the distribution itself. In [Zhang, 2004] it is

shown that the stationary density<sup>9</sup> for the variance process  $\nu$  has the form

$$p(\nu) = \left(\frac{2\kappa}{\sigma^2}\right)^{\frac{2\kappa\theta}{\sigma^2}} \frac{1}{\Gamma\left(\frac{2\kappa\theta}{\sigma^2}\right)} \cdot \nu^{\frac{2\kappa\theta}{\sigma^2}-1} e^{-\frac{2\kappa}{\sigma^2}\nu}.$$

Using the parameter values listed above, we find that  $2\kappa\theta/\sigma^2 - 1 = -0.6$ . So the density for  $\nu$  diverges for small values of  $\nu$  since then the term  $\nu^{\frac{2\kappa\theta}{\sigma^2}-1}$  diverges.

In summary we can conclude from the discussion in this section that for short-term forward-starting options the Heston model is likely to overemphasize low at-the-money/high volatility scenarios. Due to their payoff structure that depends much more on future dates than forward-starting options we expect similar limitations for exotics like cliquets, reverse cliquets, accumulators or Napoleons.

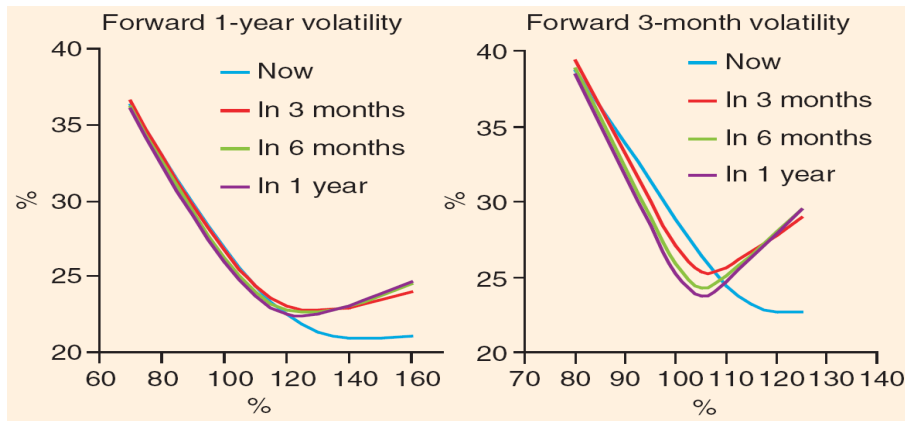


Figure 18: Forward Smile for 1-Year and 3-Month Maturity

### 3.6.6 Local Dynamics

This section deals with the local dynamics of the Heston model: How do implied volatilities move when the spot moves?

For local volatility models the following relationship between the skew and the dynamics of the at-the-money volatility as a function of the spot is well-known (see [Bergomi, 2004]):

$$\frac{d\hat{\sigma}_{K=S}}{d \ln S} = 2 \left. \frac{d\hat{\sigma}}{d \ln K} \right|_{K=S},$$

which shows that  $\hat{\sigma}_{K=S}$  moves twice as fast as the skew.

When we use the relations (3.6.1), (3.6.2), (3.6.1) and (3.6.2) we are able to derive expressions for the expected variation in the at-the-money forward volatility  $\hat{\sigma}_F$  where we have to keep in mind that conditional on a small move of the spot  $\delta S$ , the variance  $\nu$  moves on average by

$$\delta \nu = \frac{\rho \sigma}{S} \delta S. \quad (*)$$

<sup>9</sup>By the stationary density we understand the limit  $\lim_{s \rightarrow \infty} p(\nu(s)|\nu(t))$ , i.e. for large times.

**Case  $T \ll \tau$ :**

We have that

$$\hat{\sigma}_F = \sqrt{\nu} \quad \text{and} \quad \left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\rho\sigma}{4\sqrt{\nu}}.$$

This is correct at order zero in  $T$ .

The expression for the at-the-money volatility then gives:

$$[\delta\hat{\sigma}_{K=F}] = \delta\sqrt{\nu} \stackrel{(\diamond)}{=} \frac{1}{2\sqrt{\nu}}\delta\nu \stackrel{(*)}{=} \frac{\rho\sigma}{2\sqrt{\nu}} \frac{\delta S}{S}.$$

Thereby  $(\diamond)$  holds due to the relation

$$\frac{\delta\sqrt{\nu}}{\delta\nu} = \frac{1}{2\sqrt{\nu}} \quad \Leftrightarrow \quad \delta\sqrt{\nu} = \frac{1}{2\sqrt{\nu}}\delta\nu.$$

Looking at the expression for the skew we get:

$$-\frac{[\delta\hat{\sigma}_{K=F}]}{\delta \ln S} = \frac{\rho\sigma}{2\sqrt{\nu}} \cdot \frac{\delta S}{S} \cdot \frac{1}{\delta \ln S} \stackrel{(\diamond\diamond)}{=} \frac{\rho\sigma}{2\sqrt{\nu}} = 2 \left. \frac{d\hat{\sigma}}{d \ln K} \right|_{K=F}.$$

Thereby  $(\diamond\diamond)$  holds due to the relation

$$\frac{\delta \ln S}{\delta S} = \frac{1}{S} \quad \Leftrightarrow \quad \delta \ln S = \frac{1}{S}\delta S.$$

This shows that locally and for short maturities the change in implied volatilities expected by the Heston model when the spot moves is identical to that of a local volatility model, where the volatility was assumed as a deterministic function. So the skew is deterministic too which is of course not a desired property of a stochastic volatility model.

**Case  $T \gg \tau$ :**

We have that

$$\hat{\sigma}_F = \sqrt{\theta} \left( 1 + \frac{\rho\sigma}{4\kappa} \right) + \frac{\sqrt{\theta}}{2\kappa T} \left( \frac{\nu - \theta}{\theta} + \frac{\rho\sigma}{4\kappa} \frac{\nu - 3\theta}{\theta} \right) \quad \text{and} \quad \left. \frac{d\hat{\sigma}}{d \ln K} \right|_F = \frac{\rho\sigma}{2\kappa T \sqrt{\theta}}.$$

This is correct at order one in  $\frac{1}{T}$ .

Here we get by using the term for the at-the-money volatility and keeping only terms linear in  $\sigma$ :

$$\begin{aligned} [\delta\hat{\sigma}_{K=F}] &= \left[ \delta \left\{ \underbrace{\sqrt{\theta} \left( 1 + \frac{\rho\sigma}{4\kappa} \right)}_{\text{static}} + \frac{\nu}{2\kappa T \sqrt{\theta}} - \underbrace{\frac{\sqrt{\theta}}{2\kappa T}}_{\text{static}} + \frac{\rho\sigma\nu}{8\kappa^2 T \sqrt{\theta}} - \underbrace{\frac{3\sqrt{\theta}}{2\kappa T}}_{\text{static}} \right\} \right] \\ &\stackrel{(*)}{=} \frac{\rho\sigma}{2\kappa T \sqrt{\theta}} \cdot \frac{\delta S}{S} + \frac{(\rho\sigma)^2}{8\kappa^2 T \sqrt{\theta}} \cdot \frac{\delta S}{S} = \frac{\rho\sigma}{2\kappa T \sqrt{\theta}} \cdot \frac{\delta S}{S}. \end{aligned}$$

Comparing this with the expression for the skew we see that

$$-\frac{[\delta\hat{\sigma}_{K=F}]}{\delta \ln S} = \frac{\rho\sigma}{2\kappa T \sqrt{\theta}} \cdot \frac{\delta S}{S} \cdot \frac{1}{\delta \ln S} \stackrel{(\diamond\diamond)}{=} \frac{\rho\sigma}{2\kappa T \sqrt{\theta}} = \left. \frac{d\hat{\sigma}}{d \ln K} \right|_{K=F}.$$



This means that the at-the-money forward volatility slides along the smile and the Heston model behaves like a sticky-strike model: implied volatilities do not move as the spot moves. So for longer maturities the skew vanishes and which means that the Heston model cannot give the time evolution of the implied volatilities a realistic behaviour.

Therefore we need a model that can drive the dynamics of implied volatilities more realistic! In the next section we will develop the Bergomi forward variance model and show that this model is able to guarantee this.

## 4 The Bergomi Forward Variance Model

As already explained in section 3.6, a common feature of recent exotic options such as Napoleons and reverse cliquets is that their price depends not only on assumptions made for the dynamics of the underlying but also on the dynamics of implied volatilities. In other words: Their price depends on the *joint dynamics* of the underlying and its implied volatilities. These joint dynamics fall into three categories:

- The spot/volatility correlation (compare section 3.6.3).
- The dynamics of implied volatilities and more specially the term structure of the volatility of volatility (compare section 3.6.4).
- The forward skew (compare 3.6.5).

In section 3.6 we also already pointed out that although the Heston model produces prices that include an estimation of these three effects, it imposes *structural constraints* on how these features of the joint dynamics of the spot and the implied volatilities are related.

Another drawback of the Heston model (and stochastic volatility models in general) is that it is only based on the specification of the spot process and that it does not take into account that variance swaps can be considered as hedge instruments too and be endowed with their own dynamics.

In this chapter we will introduce the Bergomi forward variance model which is a consequence of these structural limits analysed in section 3.6 and the drawbacks mentioned above. The Bergomi model is based on a two-step specification:

- Set up dynamics for (discrete) forward variances.
- Set up dynamics of the spot consistent with that of the forward variances.

This finally allows us after having developed the full model specification to independently set requirements for:

- The dynamics of variance swap volatilities.
- The level of the short-term forward skew.
- The correlation between the underlying and short and long volatilities.

The Bergomi model is a *stochastic local volatility model*, it aims at the pricing of both standard exotic options and general options on variance in a consistent way and is called **fourth generation model**.

This chapter is organized as follows: First, we set up a general framework for the dynamics of forward variance swap variances. Then we specify dynamics of the underlying which are consistent with the dynamics of forward variances and derive a partial differential equation which can be used for option pricing. In the next section we specify a particular choice for the dynamics of forward variances and the underlying. We then focus on practical features of the model such as the term structure of the volatility of volatility and the term structure of the skew. Then a section focuses on pricing examples. We closely follow [Bergomi, 2005], [Bergomi, 2006] and [Korn, 2006].

## 4.1 Dynamics for the Forward Variances

In this section we specify the dynamics for the forward variances. At first, we define forward variances corresponding to a variance swap and give a special property of these forward variances. Then we introduce a one-factor model followed by a two-factor model and a discrete LIBOR-type setting and we finish with an N-factor model.

We consider now a variance swap which pays at maturity  $T$  the amount

$$V_{tT}^h - V_t^T, \quad (4.1.1)$$

where

$$\begin{aligned} V_{tT}^h &= \text{annualized variance of the spot realized over the interval } [t, T] \\ V_t^T &= \text{implied variance swap variance observed at time } t \text{ for maturity } T. \end{aligned}$$

Due to [Bergomi, 2005], variance swaps of this form are statically replicable by vanilla options. Therefore  $V_t^T$  only depends on the implied volatilities seen at time  $t$  for maturity  $T$ . Because of the definition of  $V_t^T$ , the variance swap contract has the value zero at the beginning. Since we want to specify dynamics for forward variances, we first have to define them.

### Definition 4.1 (Forward Variance Swap Variance):

Let  $T_1 < T_2$  be two maturities and  $V_t^{T_1}, V_t^{T_2}$  the corresponding implied variance swap variances at time  $t < T_1, T_2$ . Then the forward variance swap variance  $V_t^{T_1, T_2}$  is defined as

$$V_t^{T_1, T_2} := \frac{(T_2 - t)V_t^{T_2} - (T_1 - t)V_t^{T_1}}{T_2 - T_1}. \quad (4.1.2)$$

In order to model the forward variances correctly we have to factor in their properties which result from their definition.

### Proposition 4.2.

The pricing drift of any forward variance  $V_t^{T_1, T_2}$  is zero.

#### Proof:

Let us look at the costs of entering a trade whose payout at time  $t+dt$  is linear in  $V_{t+dt}^{T_1, T_2} - V_t^{T_1, T_2}$ . Therefore we consider the following trading strategy where we assume zero trading costs:

- Buy an amount  $\frac{T_2-t}{T_2-T_1} e^{r(T_2-t)}$  of a variance swap with maturity  $T_2$ .
- Sell an amount  $\frac{T_1-t}{T_2-T_1} e^{r(T_1-t)}$  of a variance swap with maturity  $T_1$ .

Then our profit and loss at time  $t' = t + dt$  reads:

$$\begin{aligned} P\&L &= \frac{T_2-t}{T_2-T_1} e^{r(T_2-t)} \cdot \left\{ \frac{V_{tt'}^h(t'-t) + V_{t'}^{T_2}(T_2-t')}{T_2-t} - V_t^{T_2} \right\} \cdot e^{-r(T_2-t')} \\ &\quad - \frac{T_1-t}{T_2-T_1} e^{r(T_1-t)} \cdot \left\{ \frac{V_{tt'}^h(t'-t) + V_{t'}^{T_1}(T_1-t')}{T_1-t} - V_t^{T_1} \right\} \cdot e^{-r(T_1-t')} \\ &= e^{r(t'-t)} \left\{ \underbrace{\frac{(T_2-t')V_{t'}^{T_2} - (T_1-t')V_{t'}^{T_1}}{T_2-T_1}}_{\stackrel{\text{Def.}}{=} V_{t'}^{T_1, T_2}} - \underbrace{\frac{(T_2-t)V_t^{T_2} - (T_1-t)V_t^{T_1}}{T_2-T_1}}_{\stackrel{\text{Def.}}{=} V_t^{T_1, T_2}} \right\} \\ &= \left\{ V_{t'}^{T_1, T_2} - V_t^{T_1, T_2} \right\} \cdot e^{r(t'-t)} = \left\{ V_{t'}^{T_1, T_2} - V_t^{T_1, T_2} \right\} r dt. \end{aligned}$$

This means that this trading strategy generates a  $P\&L$  which is linear in  $V_{t+dt}^{T_1, T_2} - V_t^{T_1, T_2}$  at lowest order in  $dt$  at zero initial costs. Thus the pricing drift of any forward variance  $V_t^{T_1, T_2}$  is zero. This completes the proof. ■

Now we are able to specify the dynamics for the variance swap curve. Therefore define the value of the variance for the date  $T$  observed at time  $t$  as

$$\xi^T(t) := V_t^{T, T}. \quad (4.1.3)$$

#### 4.1.1 A One-factor Model

Due to proposition 4.2 we are free to specify any dynamics for the  $\xi^T(t)$  that fullfill the requirement of the driftlessness. However, for practical purposes, we want to drive the dynamics of all  $\xi^T(t)$  with as small factors as possible. In the following we show how we can reach this by a suitable choice of the volatility function of  $\xi^T(t)$ .

Assume that  $\xi^T(t)$  is lognormally distributed and its volatility is a function of  $T - t$  ( $\rightarrow$  the model is invariant under time translation). So we write the following SDE:

$$d\xi^T(t) = \omega(T - t)\xi^T(t)dU(t), \quad (4.1.4)$$

where  $\{U(t)\}_{t \geq 0}$  is a Brownian motion.

But this specification would have one big drawback: For a general volatility function  $\omega(\cdot)$  the generation of  $\xi^T(t)$  for all  $T$  requires the knowledge of the complete path of  $\{U(t)\}_{t \geq 0}$  and so the model would become non-Markovian. To avoid this, we choose

$$\omega(T - t) = \omega \cdot e^{-k_1(T-t)}, \quad (4.1.5)$$

i.e.

$$d\xi^T(t) = \omega e^{-k_1(T-t)}\xi^T(t)dU(t). \quad (4.1.6)$$

Applying variation of constants the solution of this equation reads:

$$\begin{aligned} \xi^T(t) &= \xi^T(0) \exp \left( \int_0^t \omega e^{-k_1(T-u)} dU(u) - 0.5 \int_0^t \omega^2 e^{-2k_1(T-u)} du \right) \\ &= \xi^T(0) \exp \left( \omega e^{-k_1(T-t)} \int_0^t e^{-k_1(t-u)} dU(u) - 0.5 \omega^2 e^{-2k_1(T-t)} \int_0^t e^{-2k_1(t-u)} du \right) \end{aligned}$$

Defining

$$X(t) := \int_0^t e^{-k_1(t-u)} dU(u)$$

we get:

$$\xi^T(t) = \xi^T(0) \exp \left( \omega e^{-k_1(T-t)} X(t) - \frac{\omega^2}{2} e^{-2k_1(T-t)} \int_0^t e^{-2k_1(t-u)} du \right)$$

**Remark.** The process  $X(t)$  represents an Ornstein-Uhlenbeck process, since it is the solution of the following SDE:

$$dX(t) = -k_1 X(t) dt + dU(t) \quad (4.1.7)$$

$$X(0) = 0. \quad (4.1.8)$$

**Proof:**

To see this, remember the well-known variation of constants theorem which guarantees for square-integrable progressively measurable real-valued processes  $A, a, B$  and  $b$  that the linear SDE

$$dY(t) = (A(t)Y(t) + a(t))dt + (B(t)Y(t) + b(t))dW(t) \quad (*)$$

possesses the P-unique solution

$$Y(t) = Z(t) \left[ Y(0) + \int_0^t \frac{1}{Z(u)} (a(u) - B(u)b(u)) du + \int_0^t \frac{b(u)}{Z(u)} dW(u) \right],$$

where

$$Z(t) = Z(0)e^{\int_0^t (A(u) - 0.5B(u)^2) du + \int_0^t B(u)dW(u)}$$

is the solution of the homogeneous SDE

$$dZ(t) = Z(t) [A(t)dt + B(t)dW(t)].$$

Comparing the SDEs (4.1.7) and (\*), we see that

$$A(t) = -k_1, \quad a(t) = 0, \quad B(t) = 0 \quad \text{and} \quad b(t) = 1.$$

*First step:* Solve the homogeneous SDE

$$dZ(t) = -k_1 Z(t)dt \quad ; \quad Z(0) = 1.$$

This is an ordinary differential equation and has the solution

$$Z(t) = e^{-k_1 t}.$$

*Second step:* Due to the variation of constants theorem the linear SDE has then the solution

$$\begin{aligned} Y(t) &= e^{-k_1 t} \left[ \underbrace{Y(0)}_{=0} + \int_0^t \frac{1}{e^{-k_1 u}} dW(u) \right] \\ &= e^{-k_1 t} + \int_0^t e^{k_1 u} dW(u) = \int_0^t e^{-k_1(t-u)} dW(u). \end{aligned}$$

This shows the claim. ■

Now we want to connect the second integral in the expression for  $\xi^T(t)$  with the variance of the Ornstein Uhlenbeck process defined above. Therefore we first need to state a well-known result whose proof can for example be found in [Kraft, 2005].

**Lemma 4.3:**

Let  $f : [0, t] \rightarrow \mathbb{R}$  be a measurable deterministic function such that  $\int_0^t f(u)^2 du < \infty$ . Then we have:

$$\int_0^t f(u) dW(u) \sim \mathcal{N} \left( 0, \int_0^t f(u)^2 du \right).$$

Applying this to  $X(t) = \int_0^t e^{-k_1(t-u)} dU(u)$  yields:

$$\text{Var}[X(t)^2] = [X(t)^2] = \int_0^t e^{-2k_1(t-u)} du.$$

Using this we can finally write for  $\xi^T(t)$ :

$$\xi^T(t) = \xi^T(0) \exp \left( \omega e^{-k_1(T-t)} X(t) - \frac{\omega^2}{2} e^{-2k_1(T-t)} [X(t)^2] \right). \quad (4.1.9)$$

Note thereby that  $\xi^T(t)$  is driftless by construction.

Now that we know  $X(t)$  we can generate  $X(t + \delta)$ . We have that

$$\begin{aligned} X(t + \delta) &\stackrel{\text{Def.}}{=} \int_0^{t+\delta} e^{-k_1(t+\delta-u)} dU(u) = e^{-k_1\delta} \int_0^{t+\delta} e^{-k_1(t-u)} dU(u) \\ &= e^{-k_1\delta} \int_0^t e^{-k_1(t-u)} dU(u) + e^{-k_1\delta} \int_t^{t+\delta} e^{-k_1(t-u)} dU(u) \\ &\stackrel{\text{Def.}}{=} e^{-k_1\delta} X(t) + \underbrace{\int_t^{t+\delta} e^{-k_1(t+\delta-u)} dU(u)}_{=:x(\delta)} \end{aligned}$$

From lemma 4.3 we know that

$$x(\delta) \sim \mathcal{N} \left( 0, \int_t^{t+\delta} e^{-2k_1(t+\delta-u)} du \right).$$

So the variance of the integral  $x(\delta)$  is now deterministic and we can calculate its value:

$$\begin{aligned} [x(\delta)^2] &= \text{Var}[x(\delta)] = \int_t^{t+\delta} e^{-2k_1(t+\delta-u)} du = \left[ \frac{1}{2k_1} e^{-2k_1(t+\delta-u)} \right]_t^{t+\delta} \\ &= \frac{1}{2k_1} \left( \underbrace{e^{-2k_1 \cdot 0}}_{=1} - e^{-2k_1\delta} \right) = \frac{1 - e^{-2k_1\delta}}{2k_1}. \end{aligned}$$

This means for simulation purposes that we can generate  $X(t + \delta)$  using the relation

$$X(t + \delta) = e^{-k_1\delta} X(t) + x(\delta),$$

where  $x(\delta)$  is a centered Gaussian random variable such that

$$[x(\delta)^2] = \frac{1 - e^{-2k_1\delta}}{2k_1}.$$

Starting from known values for  $X(t)$  and  $[X(t)^2]$  at time  $t$  we can then generate the **one-factor forward variance curve**  $\xi^T(t + \delta)$  at time  $t + \delta$  using the relations

$$\begin{aligned} X(t + \delta) &= e^{-k_1\delta} X(t) + x(\delta) \\ [X(t + \delta)^2] &= e^{-2k_1\delta} [X(t)^2] + \frac{1 - e^{-2k_1\delta}}{2k_1} \end{aligned}$$

and

$$\xi^T(t + \delta) = \xi^T(0) \exp \left( \omega e^{-k_1(T-(t+\delta))} X(t + \delta) - \frac{\omega^2}{2} e^{-2k_1(T-(t+\delta))} [X(t + \delta)^2] \right).$$

This shows that the model has become Markovian by choosing  $\omega(T - t) = w \cdot e^{-k_1(T-t)}$  since all  $\xi^T(t)$  are functions of just one Gaussian factor  $X(t)$ .

### 4.1.2 A Two-factor Model

As already mentioned we want to drive the dynamics of the forward variances with a small number of factors. But the one-factor model is not flexible enough for practical purposes. Therefore we add a second factor in order to achieve more flexibility in the range of the term structure of the volatilities of volatilities that can be generated. This property is crucial for the model. Just remember the structural limits of the Heston model where we have shown that the volatility of volatility is overestimated roughly by a factor of two. To do this, we write for the dynamics:

$$d\xi^T(t) = \omega \xi^T(t) \left( e^{-k_1(T-t)} dU(t) + \theta e^{-k_2(T-t)} dW(t) \right), \quad (4.1.10)$$

where  $\{U(t)\}_{t \geq 0}$  and  $\{W(t)\}_{t \geq 0}$  are Brownian motions with  $\text{Corr}(U(t), W(t)) = \rho$ . Variation of constants applied to (4.1.10) yields:

$$\begin{aligned} \xi^T(t) &= \xi^T(0) \exp \left( \omega \left\{ \int_0^t e^{-k_1(T-u)} dU(u) + \theta \int_0^t e^{-k_2(T-u)} dW(u) \right\} \right. \\ &\quad \left. - \frac{\omega^2}{2} \left\{ \int_0^t e^{-2k_1(T-u)} du + \theta^2 \int_0^t e^{-2k_2(T-u)} du + 2\theta\rho \int_0^t e^{-(k_1+k_2)(T-u)} du \right\} \right) \\ &= \xi^T(0) \exp \left( \omega \left\{ e^{-k_1(T-t)} \int_0^t e^{-k_1(t-u)} dU(u) + \theta e^{-k_2(T-t)} \int_0^t e^{-k_2(t-u)} dW(u) \right\} \right. \\ &\quad \left. - \frac{\omega^2}{2} \left\{ \int_0^t e^{-2k_1(T-u)} du + \theta^2 \int_0^t e^{-2k_2(T-u)} du + 2\theta\rho \int_0^t e^{-(k_1+k_2)(T-u)} du \right\} \right) \\ &= \xi^T(0) \exp \left( \omega \left\{ e^{-k_1(T-t)} X(t) + \theta e^{-k_2(T-t)} Y(t) \right\} \right. \\ &\quad \left. - \frac{\omega^2}{2} \left\{ e^{-2k_1(T-t)} \int_0^t e^{-2k_1(t-u)} du + \theta^2 e^{-2k_2(T-t)} \int_0^t e^{-2k_2(t-u)} du \right. \right. \\ &\quad \left. \left. + 2\theta\rho e^{-(k_1+k_2)(T-t)} \int_0^t e^{-(k_1+k_2)(t-u)} du \right\} \right), \end{aligned}$$

where

$$X(t) = \int_0^t e^{-k_1(t-u)} dU(u) \quad \text{and} \quad Y(t) = \int_0^t e^{-k_2(t-u)} dW(u)$$

are Ornstein-Uhlenbeck processes.

From lemma 4.3 we deduce analogously to the one-factor case that

$$\int_0^t e^{-2k_1(t-u)} du = [X(t)^2] \quad \text{and} \quad \int_0^t e^{-2k_2(t-u)} du = [Y(t)^2].$$

To be able to rewrite the integral

$$\rho \int_0^t e^{-(k_1+k_2)(t-u)} du$$

we state without proof a further standard result similar to lemma 4.3:

**Lemma 4.4:**

Let  $f : [0, T] \rightarrow \mathbb{R}$  and  $g : [0, T] \rightarrow \mathbb{R}$  be measurable deterministic functions such that

$\int_0^t f(u)^2 du < \infty$  and  $\int_0^t g(u)^2 du < \infty$  and let  $W_1$  and  $W_2$  be two Brownian motions with  $\text{Corr}(W_1, W_2) = \rho$ . Then we have:

$$\left[ \int_0^t f(u) dW_1(u) \right] = \left[ \int_0^t g(u) dW_2(u) \right] = 0$$

and

$$\text{Cov} \left( \int_0^t f(u) dW_1(u), \int_0^t g(u) dW_2(u) \right) = \rho \int_0^t f(u) g(u) du.$$

Applying lemma 4.4 to  $X(t) = \int_0^t e^{-k_1(t-u)} dU(u)$  and  $Y(t) = \int_0^t e^{-k_2(t-u)} dW(u)$  with  $\text{Corr}(U, W) = \rho$  yields:

$$[X(t)Y(t)] = \text{Cov}(X(t), Y(t)) = \rho \int_0^t e^{-k_1(t-u)} e^{-k_2(t-u)} du = \rho \int_0^t e^{-(k_1+k_2)(t-u)} du.$$

Using these results we can finally write for the forward variances:

$$\begin{aligned} \xi^T(t) = \xi^T(0) \exp & \left( \omega \left[ e^{-k_1(T-t)} X(t) + \theta e^{-k_2(T-t)} Y(t) \right] \right. \\ & - \frac{\omega^2}{2} \left[ e^{-2k_1(T-t)} [X(t)^2] + \theta^2 e^{-2k_2(T-t)} [Y(t)^2] \right. \\ & \left. \left. + 2\theta e^{-(k_1+k_2)(T-t)} [X(t)Y(t)] \right] \right). \end{aligned} \quad (4.1.11)$$

As next we want to derive how the forward variance curve can be generated using the two-factor model. The procedure is done analogously to the one-factor case:

The same argumentation as in the one-factor case yields that

$$X(t+\delta) = e^{-k_1\delta} X(t) + x(\delta) \quad \text{and} \quad Y(t+\delta) = e^{-k_2\delta} Y(t) + y(\delta), \quad (4.1.12)$$

where

$$x(\delta) = \int_t^{t+\delta} e^{-k_1(t+\delta-u)} dU(u) \quad \text{and} \quad y(\delta) = \int_t^{t+\delta} e^{-k_2(t+\delta-u)} dW(u)$$

with

$$[x(\delta)^2] = \frac{1 - e^{-2k_1\delta}}{2k_1} \quad \text{and} \quad [y(\delta)^2] = \frac{1 - e^{-2k_2\delta}}{2k_2}. \quad (4.1.13)$$

This means for simulation purposes that we can generate  $X(t+\delta)$  and  $Y(t+\delta)$  using the relation (4.1.12), where  $x(\delta)$  and  $y(\delta)$  are Gaussian centered random variables with the variances (4.1.13). From (4.1.12) and (4.1.13) we can further deduce that

$$[X(t+\delta)^2] = e^{-2k_1\delta} [X(t)^2] + \frac{1 - e^{-2k_1\delta}}{2k_1} \quad (4.1.14)$$



and

$$[Y(t + \delta)^2] = e^{-2k_2\delta} [Y(t)^2] + \frac{1 - e^{-2k_2\delta}}{2k_2}. \quad (4.1.15)$$

The last building block that we need for the generation of the forward variance curve is  $[X(t + \delta)Y(t + \delta)]$ . Using (4.1.12) we get:

$$\begin{aligned} [X(t + \delta)Y(t + \delta)] &= [(e^{-k_1\delta}X(t) + x(\delta))(e^{-k_2\delta}Y(t) + y(\delta))] \\ &= e^{-(k_1+k_2)\delta} [X(t)Y(t)] + e^{-k_1\delta} [X(t)y(\delta)] + e^{-k_2\delta} [Y(t)x(\delta)] + [x(\delta)y(\delta)] \\ &= e^{-(k_1+k_2)\delta} [X(t)Y(t)] + [x(\delta)y(\delta)]. \end{aligned}$$

Applying lemma (4.4) we see keeping in mind the definitions of  $x(\delta)$  and  $y(\delta)$  that

$$\begin{aligned} [x(\delta)y(\delta)] &= \text{Cov}(x(\delta), y(\delta)) = \rho \int_t^{t+\delta} e^{-(k_1+k_2)(t+\delta-u)} du = \rho \left[ \frac{1}{k_1+k_2} e^{-(k_1+k_2)(t+\delta-u)} \right]_t^{t+\delta} \\ &= \frac{\rho}{k_1+k_2} \left( \underbrace{e^{-(k_1+k_2) \cdot 0}}_{=0} - e^{-(k_1+k_2)t} \right) = \rho \frac{1 - e^{-(k_1+k_2)t}}{k_1+k_2}. \end{aligned}$$

So we have:

$$[X(t + \delta)Y(t + \delta)] = e^{-(k_1+k_2)\delta} [X(t)Y(t)] + \rho \frac{1 - e^{-(k_1+k_2)t}}{k_1+k_2}. \quad (4.1.16)$$

Summarizing the calculations above we see that starting from known values  $X(t), Y(t), [X(t)^2], [Y(t)^2]$  and  $[X(t)Y(t)]$  at time  $t$  we can generate the **two-factor forward variance curve**  $\xi^T(t + \delta)$  at time  $t + \delta$  using the relations (4.1.12), (4.1.14), (4.1.14), (4.1.16) and

$$\begin{aligned} \xi^T(t + \delta) &= \xi^T(0) \exp \left( \omega \left[ e^{-k_1(T-(t+\delta))} X(t + \delta) + \theta e^{-k_2(T-(t+\delta))} Y(t + \delta) \right] \right. \\ &\quad \left. - \frac{\omega^2}{2} \left[ e^{-2k_1(T-(t+\delta))} [X(t + \delta)^2] + \theta^2 e^{-2k_2(T-(t+\delta))} [Y(t + \delta)^2] \right. \right. \\ &\quad \left. \left. + 2\theta e^{-(k_1+k_2)(T-(t+\delta))} [X(t + \delta)Y(t + \delta)] \right] \right). \end{aligned}$$

This means that starting from  $t = 0$  we can then generate a forward variance curve at any future date  $t$  by simulating two Gaussian factors and the forward variance curve is calibrated by construction to the initial forward variance curve.

We choose  $k_1 > k_2$  and call  $X(t)$  the **short factor** and  $Y(t)$  the **long factor**.

#### 4.1.3 A discrete LIBOR-type Structure

The Bergomi forward variance model aims to price recent exotic options in a consistent way. Corresponding to these exotics we have a given time interval of interest (the life time/payment relevant time of the option) and a discretization adapted to the payment time of the exotic option to value. Therefore it is instructive to set up a discrete tenor structure and model the dynamics of the forward variances for discrete time intervals in a way which is analogous to LIBOR market models instead of modelling the continuous set of all instantaneous forward

variances. This choice of a discrete tenor structure then gives us the possibility to control the skew for a given time scale.

So we assume that we have a discrete set of  $N$  equidistant dates

$$\{T_i\}_{i=0,\dots,N-1} = \{t_0 + i\Delta\}_{i=0,\dots,N-1} \quad (4.1.17)$$

starting from today's date  $t_0$ .

We will model the dynamics of the forward variances defined over the intervals of width  $\Delta$ . Therefore define

$$\xi^i(t) := V_t^{t_0+i\Delta, t_0+(i+1)\Delta} \quad \text{with} \quad , t \leq t_0 + i\Delta = T_i \quad \text{for} \quad i = 0, \dots, N-1.$$

This means that  $\xi^i(t)$  is thereby a random process until  $t = t_0 + i\Delta$ . As soon as  $t$  reaches  $t = t_0 + i\Delta$ , the variance swap variance for the time interval  $[t, t + \Delta]$  is known and is equal to

$$\xi^i(t = t_0 + i\Delta) = \xi^i(t = T_i) = V_{t_0+i\Delta}^{t_0+i\Delta, t_0+(i+1)\Delta}.$$

Figure 19 gives a graphical representation of the discrete tenor structure of the forward variances.

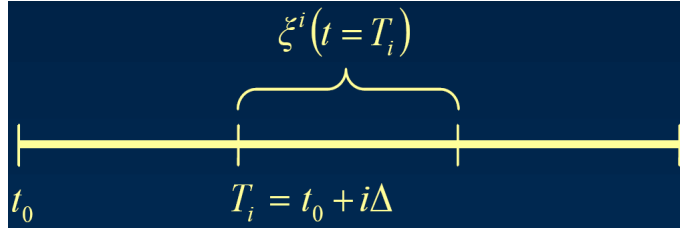


Figure 19: Tenor Structure of Forward Variances

We model the dynamics of the discrete forward variances  $\xi^i$  completely analogously to their continuous two-factor counterparts, i.e. we write the stochastic differential equations

$$d\xi^i(t) = \omega \left( e^{-k_1(T_i-t)} dU(t) + \theta e^{-k_2(T_i-t)} dW(t) \right) \quad \text{with} \quad \text{Corr}(U(t), W(t)) = \rho, \quad (4.1.18)$$

which have the solutions

$$\begin{aligned} \xi^i(t) = \xi^i(0) \exp \Big( & \omega \left\{ e^{-k_1(T_i-t)} X(t) + \theta e^{-k_2(T_i-t)} Y(t) \right\} \\ & - \frac{\omega^2}{2} \left\{ e^{-2k_1(T_i-t)} \int_0^t e^{-2k_1(t-u)} du + \theta^2 e^{-2k_2(T_i-t)} \int_0^t e^{-2k_2(t-u)} du \right. \\ & \left. + 2\theta\rho e^{-(k_1+k_2)(T_i-t)} \int_0^t e^{-(k_1+k_2)(t-u)} du \right\} \Big) \end{aligned}$$

and these can be written as

$$\begin{aligned} \xi^i(t) = \xi^i(0) \exp \Big( & \omega \left[ e^{-k_1(T_i-t)} X(t) + \theta e^{-k_2(T_i-t)} Y(t) \right] \\ & - \frac{\omega^2}{2} \left[ e^{-2k_1(T_i-t)} [X(t)^2] + \theta^2 e^{-2k_2(T_i-t)} [Y(t)^2] \right. \\ & \left. + 2\theta e^{-(k_1+k_2)(T_i-t)} [X(t)Y(t)] \right] \Big) \quad \text{for} \quad i = 0, \dots, N-1. \end{aligned} \quad (4.1.19)$$

The only difference is that the forward variances are defined according to the discrete dates  $\{T_i\}_{i=0,\dots,N-1}$ .

For the generation of the forward variance curve we use the same recursions as in the continuous two-factor case for the Ornstein-Uhlenbeck processes  $X(t)$  and  $Y(t)$  and their corresponding variances  $[X(t)^2]$  and  $[Y(t)^2]$  resp. the covariance  $[X(t)Y(t)]$ .

**Remark.** This set-up for the dynamics of the forward variances  $\xi^i$  is reminiscent of LIBOR market models which are used for the pricing of caps, floors and swaptions whose underlyings are LIBOR rates. But a big drawback is that as yet there are no market quotes for prices of caps, floors and swaptions on forward variances. Therefore the volatilities and correlations of the  $\xi^i$  cannot be calibrated to the market (compare also the section on the calibration of the Bergomi model).

#### 4.1.4 An N-factor Model

We can also drive the  $N$  discretely modelled forward variances  $\{\xi^i\}_{i=0,\dots,N-1}$  by  $N$  different factors. We write for the dynamics the stochastic differential equations

$$d\xi^i(t) = \omega_i \xi^i(t) dZ_i(t) \quad , \quad i = 0, \dots, N-1 \quad (4.1.20)$$

which have the solutions

$$\xi^i(t) = \xi^i(0) e^{\omega_i Z_i(t) - \frac{\omega_i^2}{2} t} \quad . \quad (4.1.21)$$

Thereby we correlate the Brownian motions  $\{Z_i\}_{i=0,\dots,N-1}$  with the factors  $\text{Corr} Z_i, Z_j = \rho_{ij}$  which can be chosen at will.

In the section about pricing examples we will compare pricing results obtained using an N-factor model for which  $\omega_i = \omega$  is a constant and the correlation structure of the Brownian motions is

$$\rho(Z_i, Z_j) = \theta \rho_0 + (1 - \rho) \beta^{|j-i|} \quad ,$$

where  $\theta, \rho_0, \beta \in [0, 1]$ .

## 4.2 Joint Dynamics for the Spot Process

In this section we impose dynamics for the spot process which are consistent with that of the forward variances. At first we suppose how a continuous setting would be done but then we concentrate again on a discrete setting which can correspond to a given payoff structure of an exotic option. The section is finished by the derivation of the pricing PDE corresponding to the Bergomi model.

### 4.2.1 A continuous Setting

As a first approach we use the dynamics of the instantaneous forward variances specified in equation (4.1.11) and write lognormal dynamics for the underlying, i.e.

$$dS(t) = S(t) \left[ (r - q)dt + \sqrt{\xi^t(t)} dZ(t) \right] \quad ,$$

where  $\{Z(t)\}_{t \geq 0}$  is a Brownian motion with

$$\text{Corr}(Z(t), U(t)) = \rho_{SX} \quad \text{and} \quad \text{Corr}(Z(t), W(t)) = \rho_{SY}$$

and the forward variances are given by

$$\begin{aligned} \xi^T(t) = & \xi^T(0) \exp \left( \omega \left[ e^{-k_1(T-t)} X(t) + \theta e^{-k_2(T-t)} Y(t) \right] \right. \\ & \left. - \frac{\omega^2}{2} \left[ e^{-2k_1(T-t)} [X(t)^2] + \theta^2 e^{-2k_2(T-t)} [Y(t)^2] + 2\theta e^{-(k_1+k_2)(T-t)} [X(t)Y(t)] \right] \right) \end{aligned}$$

with

$$\text{Corr}(U(t), W(t)) = \rho.$$

$r$  is thereby the interest rate and the factor  $q$  incorporates repo costs and a dividend yield. This specification yields a two-factor stochastic volatility model. In contrast to standard one-factor stochastic volatility models it has the advantage that it is calibrated by construction to the term structure of variance swap volatilities.

We further see that in this model the level of the forward skew is determined by the correlations  $\rho_{SX}, \rho_{SY}$  and  $\rho$  and the parameters  $\omega, k_1, k_2$  and  $\theta$ . But the specification itself imposes that the skew cannot be controlled separately just like in standard stochastic volatility models. To be able to do this, we switch to a discrete structure.

#### 4.2.2 A discrete Setting

In this section we show how we can control the forward skew - or in other words, the skewness of the spot price for the time scale  $\Delta$  - by specifying the spot price dynamics corresponding to a discrete tenor structure consistent with that of the forward variances. We use the tenor structure defined in (4.1.17) and the dynamics of the discrete forward variances given by expression (4.1.19).

In section 4.1.3 we have seen that at time  $t = T_i$  the variance swap variance  $\left(\hat{\sigma}_{VS}^{(i)}\right)^2$  for the time interval  $[T_i, T_i + \Delta]$  is known. Clearly we then also know the variance swap volatility  $\hat{\sigma}_{VS}^{(i)}$  for maturity  $T_i + \Delta$ . It is given by

$$\hat{\sigma}_{VS}^{(i)} = \sqrt{\xi^i(t = T_i)}. \quad (4.2.1)$$

In order to be able to specify the spot process over the interval  $[T_i, T_i + \Delta]$ , we have to make two additional requirements:

- (1) Assume that the spot price process is homogeneous over the time interval  $[T_i, T_i + \Delta]$ , i.e. the distribution of  $\frac{S(T_i + \Delta)}{S(T_i)}$  does not depend on  $S(T_i)$ .  
The reason for this requirement is that we want to **decouple** the short forward skew and the spot/volatility correlation. Imposing this condition makes the skew of maturity  $\Delta$  independent on the spot level  $S(T_i)$ . Thus for example the prices of cliquets with period  $\Delta$  will not depend on the level of the spot/volatility correlation.
- (2) Assume that the at-the-money forward skew  $\left.\frac{d\hat{\sigma}_K}{d \ln K}\right|_F$  for maturity  $T_i + \Delta$  is a deterministic function of the variance swap volatility  $\hat{\sigma}_{VS}$  or the at-the-money forward volatility  $\hat{\sigma}_{ATMF}$ . We assume here in particular that the ATMF skew is constant proportional to  $\hat{\sigma}_{ATMF}$ .

There are many suitable processes that can fulfill these objectives. One of them is the CEV diffusion which we explained in section 2.3. Note that we need to correlate the spot process

with the process of the forward variances  $\xi^i$  for  $j > i$ .

So in particular we assume that the dynamics of the spot price read

$$dS(t) = S(t) \left[ (r(t) - q(t))dt + \sigma_0^{(i)} \left( \frac{S(t)}{S(T_i)} \right)^{1-\beta^{(i)}} dZ(t) \right] \quad \text{for } t \in [T_i, T_i + \Delta], \quad (4.2.2)$$

where the volatility parameters  $\sigma_0^{(i)}$  and  $\beta^{(i)}$  are determined by the variance swap volatility at time  $t = T_i$  for the maturity  $T_i + \Delta$ , i.e. we have

$$\sigma_0^{(i)} = \sigma_0 \left( \hat{\sigma}_{VS}^{(i)} \right) \quad \text{and} \quad \beta^{(i)} = \beta \left( \hat{\sigma}_{VS}^{(i)} \right) \quad \text{with} \quad \hat{\sigma}_{VS}^{(i)} = \sqrt{\xi^i(t = T_i)} \quad (4.2.3)$$

such that the condition (2) on the ATMF skew is fulfilled.

$r(t)$  and  $q(t)$  thereby denote the interest rate and the repo costs including the dividend yield.

For the 2-factor model the forward variances are given by

$$\begin{aligned} \xi^i(t) = \xi^i(0) \exp \bigg( & \omega \left[ e^{-k_1(T_i-t)} X(t) + \theta e^{-k_2(T_i-t)} Y(t) \right] \\ & - \frac{\omega^2}{2} \left[ e^{-2k_1(T_i-t)} [X(t)^2] + \theta^2 e^{-2k_2(T_i-t)} [Y(t)^2] \right. \\ & \left. + 2\theta e^{-(k_1+k_2)(T_i-t)} [X(t)Y(t)] \right] \bigg) \quad \text{for } i = 0, \dots, N-1 \end{aligned} \quad (4.2.4)$$

and the Brownian motions  $\{U(t)\}_{t \geq 0}$ ,  $\{W(t)\}_{t \geq 0}$  and  $\{Z(t)\}_{t \geq 0}$  are correlated as follows:

$$\text{Corr}(Z(t), U(t)) = \rho_{SX} \quad , \quad \text{Corr}(Z(t), W(t)) = \rho_{SY} \quad \text{and} \quad \text{Corr}(U(t), W(t)) = \rho. \quad (4.2.5)$$

By the equations (4.2.2), (4.2.3), (4.2.4) and (4.2.5) we have then completely specified the **Bergomi 2-factor forward variance model** according to a discrete tenor structure.

For the  $N$ -factor model the forward variances are given by

$$\xi^i(t) = \xi^i(0) e^{\omega_i Z_i(t) - \frac{\omega_i^2}{2} t} \quad \text{for } i = 0, \dots, N-1 \quad (4.2.6)$$

where the Brownian motions  $\{Z_i\}_{i=0, \dots, N-1}$  are correlated with a factor of

$$\text{Corr}(Z_i, Z_j) = \rho_{ij} \quad (4.2.7)$$

to each other and with a factor of

$$\text{Corr}(Z, Z_i) = \rho_{Si} \quad (4.2.8)$$

the Brownian motion  $Z$  of the spot process.

By the equations (4.2.2), (4.2.3), (4.2.6), (4.2.7) and (4.2.8) we have then completely specified the **Bergomi N-factor forward variance model** according to a discrete tenor structure.

From the specifications given above we obtain that the 2-factor Bergomi model as well as the  $N$ -factor Bergomi model is a *stochastic local volatility model*, since the spot process is specified as CEV process and the volatility parameters are functions of the forward variances which were specified as stochastic processes.

### 4.3 The Pricing Equation

In this section we derive partial differential equations for the Bergomi model which can be used for the pricing of exotic options. We start with the PDE corresponding to the 2-factor model followed by the PDE corresponding to the N-factor model. Our procedure is as follows:

- Assume that there exists a price process which depends on the asset price, on the forward variances and on time.
- Deduce the dynamics of this price process by applying Ito's formula.
- Assume that we are risk-free and therefore we can get rid off the terms driven by a Brownian motion.
- The terms which are left have to equal the return which is obtained for the money market account.

#### 4.3.1 The 2-factor Model

In the 2-factor model the dynamics of the forward variances  $\{\xi^i\}_{i=0,\dots,N-1}$  are driven by the Ornstein-Uhlenbeck processes  $X$  and  $Y$ .

Therefore we assume that there exists a price process

$$P = P(S; X, Y; t) \in C^{1,2} \quad (4.3.1)$$

of an option being priced such that the time  $t$ -price of the option is given by

$$C(t) = P(S(t); X(t), Y(t); t) \quad , \quad t \in [T_i, T_i + \Delta[. \quad (4.3.2)$$

Applying Ito's formula to the price process (4.3.1) yields:

$$\begin{aligned} dP = & \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} dS + \frac{\partial P}{\partial X} dX + \frac{\partial P}{\partial Y} dY + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} d\langle S \rangle_t + \frac{1}{2} \frac{\partial^2 P}{\partial X^2} d\langle X \rangle_t \\ & + \frac{1}{2} \frac{\partial^2 P}{\partial Y^2} d\langle Y \rangle_t + \frac{\partial^2 P}{\partial S \partial X} d\langle S, X \rangle_t + \frac{\partial^2 P}{\partial S \partial Y} d\langle S, Y \rangle_t + \frac{\partial^2 P}{\partial X \partial Y} d\langle X, Y \rangle_t. \end{aligned}$$

Remember thereby that the dynamics of  $S, X$  and  $Y$  read

$$\begin{aligned} dS &= (r(t) - q(t))Sdt + \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} SdZ(t), \\ dX &= -k_1 Xdt + dU(t), \\ dY &= -k_2 Ydt + dW(t), \end{aligned}$$

and they are correlated with the factors

$$\text{Corr}(Z(t), U(t)) = \rho_{SX} \quad , \quad \text{Corr}(Z(t), W(t)) = \rho_{SY} \quad \text{and} \quad \text{Corr}(U(t), W(t)) = \rho.$$

From this we can deduce the quadratic (co-) variations:

$$\begin{aligned} d\langle S \rangle_t &= \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 dt, \\ d\langle X \rangle_t &= dt, \\ d\langle Y \rangle_t &= dt, \\ d\langle S, X \rangle_t &= \rho_{SX} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S dt, \\ d\langle S, Y \rangle_t &= \rho_{SY} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S dt, \\ d\langle X, Y \rangle_t &= \rho dt. \end{aligned}$$

So we can write for the dynamics of the price process:

$$\begin{aligned} dP &= \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} \left( (r(t) - q(t)) S dt + \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S dZ(t) \right) \\ &\quad + \frac{\partial P}{\partial X} (-k_1 X dt + dU(t)) + \frac{\partial P}{\partial Y} (-k_2 Y dt + dW(t)) \\ &\quad + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 dt + \frac{\partial^2 P}{\partial S \partial X} \rho_{SX} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S dt \\ &\quad + \frac{\partial^2 P}{\partial S \partial Y} \rho_{SY} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S dt + \frac{1}{2} \frac{\partial^2 P}{\partial X^2} dt + \frac{1}{2} \frac{\partial^2 P}{\partial Y^2} dt + \frac{\partial^2 P}{\partial X \partial Y} \rho dt. \end{aligned}$$

Now **assume** that we are instantaneously risk-free. If this is the case, the terms driven by a Brownian motion vanish and due to the no-arbitrage paradigm the dynamics that are then left for the price process of  $P$  have to equal those of the money market account. So we get:

$$\begin{aligned} dP &= \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} (r(t) - q(t)) S dt - \frac{\partial P}{\partial X} k_1 X dt - \frac{\partial P}{\partial Y} k_2 Y dt \\ &\quad + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 dt + \frac{1}{2} \left( \frac{\partial^2 P}{\partial X^2} + \frac{\partial^2 P}{\partial Y^2} + 2\rho \frac{\partial^2 P}{\partial X \partial Y} \right) dt \\ &\quad + \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S \left( \rho_{SX} \frac{\partial^2 P}{\partial S \partial X} + \rho_{SY} \frac{\partial^2 P}{\partial S \partial Y} \right) dt \\ &\stackrel{!}{=} rP dt. \end{aligned}$$

After getting rid off the  $dt$  we can now finally write the **partial differential pricing equation corresponding to the 2-factor Bergomi model**:

$$\begin{aligned} \frac{\partial P}{\partial t} + (r(t) - q(t)) S \frac{\partial P}{\partial S} - k_1 X \frac{\partial P}{\partial X} - k_2 Y \frac{\partial P}{\partial Y} + \frac{1}{2} \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 \frac{\partial^2 P}{\partial S^2} \\ + \frac{1}{2} \left( \frac{\partial^2 P}{\partial X^2} + \frac{\partial^2 P}{\partial Y^2} + 2\rho \frac{\partial^2 P}{\partial X \partial Y} \right) + \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S \left( \rho_{SX} \frac{\partial^2 P}{\partial S \partial X} + \rho_{SY} \frac{\partial^2 P}{\partial S \partial Y} \right) = rP, \end{aligned} \tag{4.3.3}$$

where  $t \in [T_i, T_i + \Delta]$ .

### 4.3.2 The N-factor Model

In the N-factor model the dynamics of the forward variances  $\{\xi^i\}_{i=0,\dots,N-1}$  are driven by N factors.

Therefore we assume that there exists a price process

$$P = P(S; \xi^0, \dots, \xi^{N-1}; t) \in C^{1,2} \quad (4.3.4)$$

of an option being priced such that the time  $t$ -price of the option is given by

$$C(t) = P(S(t); \xi^0(t), \dots, \xi^{N-1}(t); t) \quad , \quad t \in [T_i, T_i + \Delta[. \quad (4.3.5)$$

Applying the multi-dimensional Ito formula to the price process (4.3.4) yields:

$$\begin{aligned} dP = & \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} dS + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} d\langle S \rangle_t + \sum_{i=0}^{N-1} \frac{\partial P}{\partial \xi^i} d\xi^i \\ & + \frac{1}{2} \sum_{i,j=0}^{N-1} \frac{\partial^2 P}{\partial \xi^i \partial \xi^j} d\langle \xi^i, \xi^j \rangle_t + \sum_{i=0}^{N-1} \frac{\partial^2 P}{\partial S \partial \xi^i} d\langle S, \xi^i \rangle_t. \end{aligned}$$

Remember here the dynamics of  $S$  and the  $\xi^i$ :

$$\begin{aligned} dS &= (r(t) - q(t))Sdt + \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} SdZ(t), \\ d\xi^i &= \omega_i \xi^i dZ_i(t). \end{aligned}$$

They are correlated with the factors

$$\text{Corr}(Z(t), Z_i(t)) = \rho_{Si} \quad \text{and} \quad \text{Corr}(Z_i(t), Z_j(t)) = \rho_{ij}.$$

This determines the quadratic (co-) variations:

$$\begin{aligned} d\langle S \rangle_t &= \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 dt, \\ d\langle \xi^i, \xi^j \rangle_t &= \rho_{ij} \omega_i \omega_j \xi^i \xi^j dt, \\ d\langle S, \xi^i \rangle_t &= \rho_{Si} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S \omega_i \xi^i dt. \end{aligned}$$

So we can write for the dynamics of the price process:

$$\begin{aligned} dP = & \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} \left( (r(t) - q(t))Sdt + \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} SdZ(t) \right) \\ & + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 dt + \sum_{i=0}^{N-1} \frac{\partial P}{\partial \xi^i} \omega_i \xi^i dZ_i(t) \\ & + \frac{1}{2} \sum_{i,j=0}^{N-1} \frac{\partial^2 P}{\partial \xi^i \partial \xi^j} \rho_{ij} \omega_i \omega_j \xi^i \xi^j dt + \sum_{i=0}^{N-1} \frac{\partial^2 P}{\partial S \partial \xi^i} \rho_{Si} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S \omega_i \xi^i dt. \end{aligned}$$



Assuming again that we are risk-free yields analogously to the 2-factor case:

$$\begin{aligned} dP &= \frac{\partial P}{\partial t} dt + \frac{\partial P}{\partial S} (r(t) - q(t)) S dt + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 dt \\ &\quad + \frac{1}{2} \sum_{i,j=0}^{N-1} \frac{\partial^2 P}{\partial \xi^i \partial \xi^j} \rho_{ij} \omega_i \omega_j \xi^i \xi^j dt + \sum_{i=0}^{N-1} \frac{\partial^2 S}{\partial S \partial \xi^i} \rho_{Si} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S \omega_i \xi^i dt \\ &\stackrel{!}{=} rP dt . \end{aligned}$$

So we can finally deduce the **partial differential pricing equation corresponding to the N-factor Bergomi model**:

$$\begin{aligned} \frac{\partial P}{\partial t} + (r(t) - q(t)) S \frac{\partial P}{\partial S} + \frac{1}{2} \left( \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} \right)^2 S^2 \frac{\partial^2 P}{\partial S^2} \\ + \frac{1}{2} \sum_{i,j=0}^{N-1} \rho_{ij} \omega_i \omega_j \xi^i \xi^j \frac{\partial^2 P}{\partial \xi^i \partial \xi^j} + \sum_{i=0}^{N-1} \rho_{Si} \sigma_0^{(i)} \left( \frac{S}{S(T_i)} \right)^{1-\beta^{(i)}} S \omega_i \xi^i \frac{\partial^2 S}{\partial S \partial \xi^i} = rP , \end{aligned} \quad (4.3.6)$$

where  $t \in [T_i, T_i + \Delta]$ .

#### How to price?

Having specified the model and having derived pricing equations, a natural question is: How can we price options using the Bergomi model?

The answer might sound very simple. Since as far there are no explicit formulas available, one has to calculate prices via Monte Carlo simulation or using a numerical partial differential equation solver applied to equation (4.3.3) resp. (4.3.6) after having calibrated the model. How the model is calibrated is the topic of the next section.

#### 4.4 Three-step Calibration

A quite important task is surely the calibration of the Bergomi model to market prices of options. In this section we show how this is done focussing on the 2-factor model. To start we always have to specify the type of options that should be priced and choose the time scale  $\Delta$  according to the payoff structure of this option type. In what follows, we assume that  $\Delta = 1$  month. Remember that by construction the model is calibrated at time  $t_0$  to the forward variance curve for all maturities  $\{T_i\}_{i=0,\dots,N-1} = \{t_0 + i\Delta\}_{i=0,\dots,N-1}$ .

In this section we will introduce a three-step procedure for the calibration:

- **Step 1:** Set dynamics for the implied variance swap volatilities, i.e. choose values for the parameters  $k_1, k_2, \omega, \rho, \theta$ .
- **Step 2:** Calibrate the short-term forward skew, i.e. determine the volatility parameters  $\sigma_0^{(i)}$  and  $\beta^{(i)}$  as functions of the implied volatility of the variance swap.
- **Step 3:** Calibrate the term skew, i.e. set the correlations  $\rho_{SX}$  and  $\rho_{SY}$  between the spot and the short-term volatility factor resp. the spot and the long-term volatility factor.

#### 4.4.1 Set Dynamics for the implied Variance Swap Volatilities

The aim of the Bergomi model is to price options whose price is a very non-linear function of the volatility (compare the graph of the price of a Napoleon option as a function of volatility in figure 12). As explained at the beginning of this chapter, we want to be able to control the term structure of the volatilities. In the Bergomi model, their dynamics are determined by variance swap volatilities and the dynamics of the variance swap volatilities are controlled by the parameters  $k_1, k_2, \omega, \rho$  and  $\theta$  in the 2-factor model.

Unfortunately, there is currently no active market for options on the variance swap volatility  $\hat{\sigma}_{VS}$ . Therefore the parameters listed above cannot be calibrated to market prices. *Thus, their values have to be chosen!*

In order to reflect the market's view in some way, these parameters have to be chosen such that the level and the term structure of the volatility of volatility are consistent with historically observed volatilities of implied volatilities. As a measure of the volatility of the variance swap volatility with maturity  $\tau$  for a given time scale  $\Delta t$  we define

$$\sigma_{Vol}(\tau) = \frac{1}{\sqrt{\Delta t}} StDev \left[ \ln \left( \frac{\sqrt{V_{\Delta t, \Delta t + \tau}}}{\sqrt{V_0}} \right) \right]. \quad (4.4.1)$$

Figure 20 gives a graphical illustration for this definition.

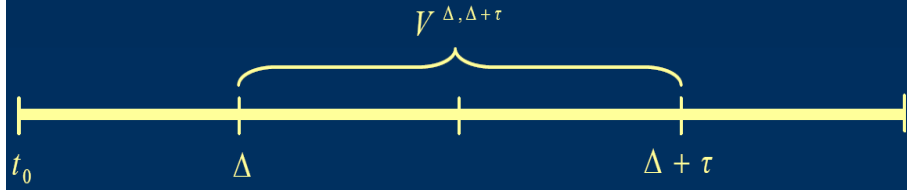


Figure 20: Illustration for the Definition of  $\sigma_{Vol}(\tau)$

We now choose the values

$$\omega = 2.827 \quad , \quad \rho = 0 \quad , \quad \theta = 30\% \quad , \quad k_1 = 6 \text{ (2 months)} \quad , \quad k_2 = 0.25 \text{ (4 years)}. \quad (4.4.2)$$

Using these values for the specification of the forward variances we get that

$$\sigma_{Vol}(1 \text{ month}) = 120\% \quad , \quad \sigma_{Vol}(1 \text{ year}) = 45\% \quad \text{and} \quad \sigma_{Vol}(5 \text{ years}) = 25\%$$

for a one-month horizon  $\Delta t = 1 \text{ month}$ . Due to [Bergomi, 2005], this is quite consistent with what one would deduce from historical values.

Figure 21 displays the complete term structure of the volatilities of the variance swap volatilities  $\{\sigma_{Vol}(\tau)\}_{\tau=1 \text{ month}, \dots, 60 \text{ months}}$  for a one-month horizon generated by the 2-factor model with a flat initial variance swap term structure at 20% volatility (i.e.  $\sqrt{\xi^i(0)} = 20\%$ ) using the values of  $k_1, k_2, \omega, \rho$  and  $\theta$  listed in (4.4.2).

For comparison we also plot the term structure generated by the N-factor model with  $\omega_i = \omega \forall i$ , i.e.

$$\xi^i(t) = \xi^i(0) e^{\omega Z_i(t) - \frac{\omega^2}{2} t}.$$

and the correlation structure

$$\rho_{ij} = \theta \rho_0 + (1 - \rho) \beta^{|j-i|}.$$

Thereby the values

$$\omega = 2.827 \quad , \quad \theta = 40\% \quad , \quad \rho_0 = 5\% \quad \text{and} \quad \beta = 10\%$$

are chosen such that the term-structure of the 2-factor model is matched.

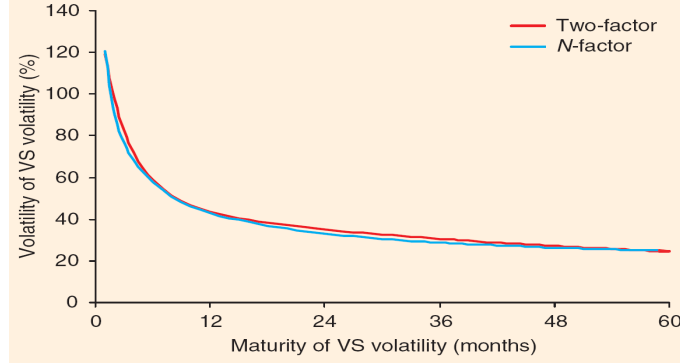


Figure 21: Term Structure of  $\sigma_{Vol}(\tau)$  for a One-month Interval

Let us now measure volatilities over a time interval  $\Delta t$  of one year. We have plotted them using the same parameters as for the one-month interval in figure 22. We see that they are very different for the 2-factor and the N-factor model. From this we can deduce that although both models would yield similar prices for options on variance swaps observed one month from now, they would price very differently options on variance swap volatilities observed in one year. An explanation for this difference is: In the 2-factor model, the volatility of volatility will tend to decrease as the time scale over which they are measured increases due to the mean-reverting nature of the driving process. In contrast to that, in the N-factor model they increase due to the fact that forward variances are log-normal.

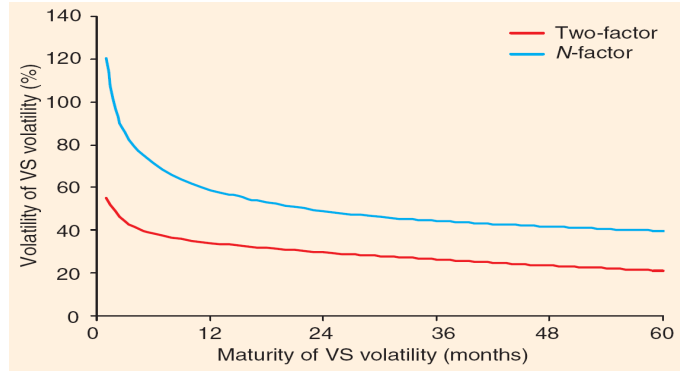


Figure 22: Term Structure of  $\sigma_{Vol}(\tau)$  for a One-year Interval

#### 4.4.2 Calibration of the short Forward Skew

In section 4.2.2, where we introduced joint dynamics for the spot, we assumed that the ATMF skew  $\frac{d\hat{\sigma}_K}{d\ln K}\Big|_F$  for maturity  $T_i + \Delta$  is

- constant or
- proportional to the ATMF volatility  $\hat{\sigma}_{ATMF}$ .

In this section we determine the volatility parameters  $\sigma_0$  and  $\beta$  as functions of the implied variance swap volatility  $\hat{\sigma}_{VS}^{(i)} = \sqrt{\xi^i(T_i)}$  for these two cases.

##### Case 1: Constant Skew.

Calibrate the functions  $\sigma_0(\hat{\sigma}_{VS}^{(i)})$  and  $\beta(\hat{\sigma}_{VS}^{(i)})$  such that the one-month ATMF skew has the constant value 5%. Here we use the 95 – 105% skew:

$$\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%} \approx -\frac{1}{10} \frac{d\hat{\sigma}_K}{d\ln K}\Big|_F \stackrel{!}{=} 5\%. \quad (4.4.3)$$

This requirement completely determines the functions  $\sigma_0(\hat{\sigma}_{VS}^{(i)})$  and  $\beta(\hat{\sigma}_{VS}^{(i)})$  and the calibration can easily be done numerically.

##### Case 2: Proportional Skew.

Calibrate the functions  $\sigma_0(\hat{\sigma}_{VS}^{(i)})$  and  $\beta(\hat{\sigma}_{VS}^{(i)})$  such that the one-month ATMF skew is proportional to the ATMF volatility such that the skew is equal to 5% when the ATMF volatility is equal to 20%. We use again the 95 – 105% skew:

$$\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%} \approx -\frac{1}{10} \frac{d\hat{\sigma}_K}{d\ln K}\Big|_F \stackrel{!}{=} 5\% \quad \text{when} \quad \hat{\sigma}_{ATMF} = 20\%. \quad (4.4.4)$$

Again this completely determines the functions  $\sigma_0(\hat{\sigma}_{VS}^{(i)})$  and  $\beta(\hat{\sigma}_{VS}^{(i)})$ .

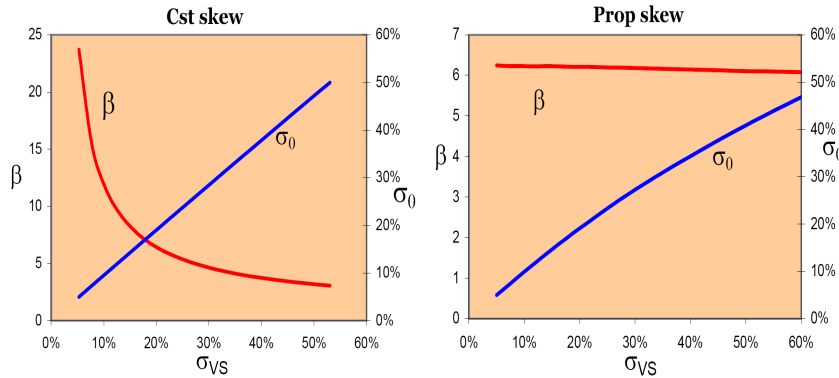


Figure 23:  $\sigma_0(\hat{\sigma}_{VS}^{(i)})$  and  $\beta(\hat{\sigma}_{VS}^{(i)})$  for constant Skew and proportional Skew

Figure 23 shows the functions  $\sigma_0(\hat{\sigma}_{VS}^{(i)})$  and  $\beta(\hat{\sigma}_{VS}^{(i)})$  for the case of a constant 95 – 105% skew equal to 5% and for the case of a proportional skew such that the 95 – 105% skew is equal to

5% when  $\hat{\sigma}_{ATMF} = 20\%$ .

In both cases, if needed, an individual calibration of  $\sigma_0 \left( \hat{\sigma}_{VS}^{(i)} \right)$  and  $\beta \left( \hat{\sigma}_{VS}^{(i)} \right)$  can be done for each time interval  $\{[T_i, T_{i+1}]\}_{i=0, \dots, N-1}$ . But typically it is sufficient for practical purposes to use the same calibration for all intervals except the first one. For the first interval a specific calibration is performed so as to match the short vanilla skew.

The level of the 95 – 105% skew can thereby be selected by the trader or chosen such that the market prices of call spread cliquets<sup>10</sup> of period  $\Delta$  (here  $\Delta = 1$  month was used) are matched.

#### 4.4.3 Setting Correlations between the Spot and short/long Factors - The Term Skew

The correlation between the spot and the short factor  $\rho_{SX}$  and the correlation between the spot and the long factor  $\rho_{SY}$  cannot be chosen independently since  $X$  and  $Y$  are itself correlated with a factor of  $\rho$ . Therefore we use the following parametrization:

$$\rho_{SY} = \rho_{SX}\rho + \chi\sqrt{(1 - \rho_{SX}^2)(1 - \rho^2)} \quad \text{with} \quad \chi \in [-1, 1]. \quad (4.4.5)$$

$\rho_{SX}$  and  $\rho_{SY}$  control both the correlation between the spot and short and long variance swap volatilities and the term structure of the skew of vanilla options. The dependence of the term skew on them is show explicitly in the following part about the term skew.

$\rho_{SX}$  and  $\rho_{SY}$  can be chosen, calibrated to the market prices of call spread cliquets of a period larger than  $\Delta$  or calibrated to the vanilla skew corresponding to the maturity of the option which is considered.

#### The Term Skew

To show how the Bergomi model generates the skew, we derive an approximate expression for the ATMF skew as a function of maturity for the case of a flat term structure of the variance swap volatilities at order one in  $\omega$  and the skew  $\left. \frac{d\hat{\sigma}_K}{d \ln K} \right|$  at time scale  $\Delta$ . We then denote this expression with  $\text{Skew}_\Delta$ . We proceed as follows:

Consider the maturity  $T = N\Delta$  and define the returns  $r_i$  via

$$r_i := \ln \left( \frac{S(i\Delta)}{F(i\Delta)} \right) - \ln \left( \frac{S((i-1)\Delta)}{F((i-1)\Delta)} \right),$$

where  $F(i\Delta)$  is the forward price for maturity  $i\Delta$ . From this definition we get that

$$\sum_{i=1}^N r_i = \ln \left( \frac{S(N\Delta)}{F(N\Delta)} \right) - \ln \left( \frac{S((0)\Delta)}{F((0)\Delta)} \right) \stackrel{F(0)=S(0)}{=} \ln \left( \frac{S(N\Delta)}{F(N\Delta)} \right) = \ln \left( \frac{S(T)}{F(T)} \right). \quad (4.4.6)$$

---

<sup>10</sup>A call spread cliquet has the payoff

$$\max \left\{ \sum_{i=1}^N R_i; \text{MaxCoupon} \right\},$$

where the  $R_i$  are call spreads of the structure

$$R_i = \max \{S(t_i) - K_1(t_i); 0\} - \max \{S(t_i) - K_2(t_i); 0\} - \alpha$$

with

$$K_1(t_i) = (1 - \alpha)S(t_{i-1}) \quad \text{and} \quad K_2(t_i) = (1 + \alpha)S(t_{i-1})$$

and a given time period  $\Delta = t_i - t_{i-1}$ , which is for example equal to one month.

In [Backus et al., 1997] it is shown that given the skewness  $\gamma\left(\ln\left(\frac{S(T)}{F(T)}\right)\right)$ , the ATMF skew at first order in  $\gamma\left(\ln\left(\frac{S(T)}{F(T)}\right)\right)$  is given by:

$$\text{Skew}_{T=N\Delta} = \frac{\gamma\left(\ln\left(\frac{S(T)}{F(T)}\right)\right)}{6\sqrt{T}}. \quad (4.4.7)$$

Remember the definition of the skewness of a random variable  $X$  with  $[X^3] < \infty$ :

$$\gamma(X) = \frac{[(X - [X])^3]}{(\text{Var}[X])^{\frac{3}{2}}}.$$

Assuming that the mean of returns is negligible, we can write for the skewness of  $\ln\left(\frac{S(T)}{F(T)}\right)$  using relation (4.4.6):

$$\gamma\left(\ln\left(\frac{S(T)}{F(T)}\right)\right) = \gamma\left(\sum_{i=1}^N r_i\right) = \frac{\left[\left(\sum_{i=1}^N r_i\right)^3\right]}{\left(\left[\left(\sum_{i=1}^N r_i\right)^2\right]\right)^{\frac{3}{2}}}.$$

So if we want to derive an expression for  $\text{Skew}_{N\Delta}$ , we have to determine the third and second moment of  $\sum_{i=1}^N r_i$ .

Since returns are uncorrelated and assuming that  $\Delta$  is small we have that

$$\left[\left(\sum_{i=1}^N r_i\right)^3\right] = \sum_{i=1}^N [r_i^3] + 3 \sum_{j>i}^N [r_i r_j^2].$$

In [Bergomi, 2005] it is then shown using the approximations

$$r_j^2 = \Delta \xi^j(T_j) \quad \text{and} \quad r_i = \sqrt{\xi^i(T_i)} \int_{T_i}^{T_i+\Delta} dZ(t)$$

that

$$\left[\left(\sum_{i=1}^N r_i\right)^3\right] = N\gamma\left(\ln\left(\frac{S(\Delta)}{F(\Delta)}\right)\right) (\xi\Delta)^{\frac{3}{2}} + \omega (\xi\Delta)^{\frac{3}{2}} N^2 (\rho_{SX}\zeta(k_1\Delta, N) + \theta\rho_{SY}\zeta(k_2\Delta, N))$$

at order one in  $\omega$  and  $S(\Delta)$ , where

$$\zeta(x, N) := \left(\frac{1 - e^{-x}}{x}\right) \frac{\sum_{\tau=1}^{N-1} (N - \tau) e^{-(\tau-1)x}}{N^2} \quad \text{and} \quad \xi = \xi^i(0) \forall i$$

and that

$$\left[\left(\sum_{i=1}^N r_i\right)^2\right] = N\xi\Delta$$

at order one in  $\omega$  and  $S(\Delta)$ . Hence we get:

$$\begin{aligned} \gamma \left( \ln \left( \frac{S(T)}{F(T)} \right) \right) &= \frac{N \gamma \left( \ln \left( \frac{S(\Delta)}{F(\Delta)} \right) \right) (\xi \Delta)^{\frac{3}{2}}}{(N \xi \Delta)^{\frac{3}{2}}} + \frac{\omega (\xi \Delta)^{\frac{3}{2}} N^2 (\rho_{SX} \zeta(k_1 \Delta, N) + \theta \rho_{SY} \zeta(k_2 \Delta, N))}{(N \xi \Delta)^{\frac{3}{2}}} \\ &= \frac{\gamma \left( \ln \left( \frac{S(\Delta)}{F(\Delta)} \right) \right)}{\sqrt{N}} + \omega \sqrt{N \Delta} (\rho_{SX} \zeta(k_1 \Delta, N) + \theta \rho_{SY} \zeta(k_2 \Delta, N)) \end{aligned}$$

Plugging this into expression (4.4.7) for the ATMF skew gives:

$$\begin{aligned} \text{Skew}_{N\Delta} &= \frac{\gamma \left( \ln \left( \frac{S(\Delta)}{F(\Delta)} \right) \right)}{6 \sqrt{N \Delta} \sqrt{N}} + \frac{\omega \sqrt{N \Delta} (\rho_{SX} \zeta(k_1 \Delta, N) + \theta \rho_{SY} \zeta(k_2 \Delta, N))}{6 \sqrt{N \Delta}} \\ &= \underbrace{\frac{\text{Skew}_{\Delta}}{N}}_{(1)} + \underbrace{\frac{\omega}{6} (\rho_{SX} \zeta(k_1 \Delta, N) + \theta \rho_{SY} \zeta(k_2 \Delta, N))}_{(2)}. \end{aligned} \quad (4.4.8)$$

Expression (4.4.8) is very instructive: It makes appearant how much of the skew at maturity  $T$  is contributed on the one hand by the intrinsic skewness of the spot process at time scale  $\Delta$  (part (1)) and on the other hand by the spot/volatility correlation (part (2)). Further we see that when  $\omega = 0$ , the skew decays as  $1/T$  which we would expect for a process with independent increments. And finally this equation shows that  $\rho_{SX}$  and  $\rho_{SY}$  can naturally be used to control the term structure of the skew when we look at expression (2).

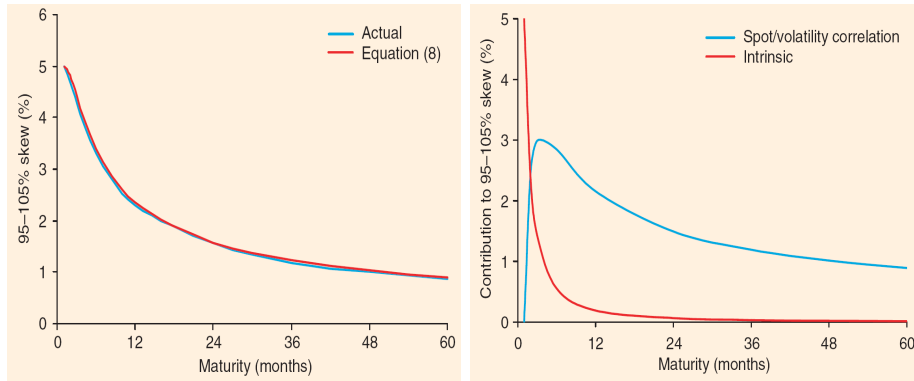


Figure 24: Left: The 95-105% Skew as a Function of Maturity compared to (4.4.8)  
Right: The two Contributions of the 95-105% Skew in (4.4.8)

The left-hand side of figure 24 shows how the approximate skew in equation (4.4.8) looks compared to the actual 95-105% skew. Thereby we have chosen  $\Delta = 1$ , month,  $(\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%})_{1 \text{ month}} = 5\%$ ,  $\omega = 2.827$ ,  $\rho = 0$ ,  $\theta = 30\%$ ,  $k_1 = 6$  and  $k_2 = 0.25$  and the spot volatility parameters are  $\rho_{SX} = -70\%$ ,  $\rho_{SY} = -35.7\%$  ( $\chi = -50\%$ ). Even though  $\omega$  and  $\text{Skew}_{\Delta}$  are both large, the actual skew is fitted very well by the approximate skew.

The two contributions to  $\text{Skew}_{N\Delta}$  in (4.4.8) are illustrated on the right-hand side of figure 24. The contribution of  $\text{Skew}_{\Delta}$  to  $\text{Skew}_{N\Delta}$  is monotonically decreasing. The contribution of the spot/volatility correlation starts from 0, then strongly increases and slowly decreases afterwards. From this behaviour we can deduce that depending on the relative magnitude of both terms, the term structure of the skew can be non-monotonic.

## 4.5 Pricing Examples

In this section we discuss how the Bergomi model prices a reverse cliquet, a Napoleon and an accumulator and we analyse the relative contribution of forward skew, volatility of volatility and spot/volatility correlation effects to prices. For this section we assume zero interest rates and dividend yield. The numerical results of this section are taken out of [Bergomi, 2005]. For the sake of comparing prices we have to specify how the model parameters are calibrated. The selection of the right instrument on which to calibrate is essential:

- When we price options that can be hedged with a static position in vanilla options, it is natural to calibrate the model to the vanilla smile (surface). But this case is rather an exception.
- When Napoleons and reverse cliquets are priced which have a high sensitivity to the forward volatility and the forward skew, it is more appropriate to calibrate the model to call spread cliquets and ATM cliquets<sup>11</sup>
- When options on the variance are priced it is natural to calibrate the model to the term structure of variance swap volatilities.

These exotic products are also very sensitive to the volatility of volatility and they are usually designed such that their price at inception is small but increases significantly if the implied volatility decreases (just compare again figure 12 which shows the price of a Napoleon as a function of volatility.)

As there is as already mentioned as yet no active market for options on variance, we use the volatility of volatility parameters listed in (4.4.2) in section 4.4.1. As long as the forward skew is not turned off, the constant 95 – 105% one-month skew is calibrated such that the price of a three-year 95 – 105% one-month call spread cliquet has a constant value, equal to its price when the volatility of volatility is turned off and the one-month 95 – 105% skew is 5%, which is equal to 191.6%.

In all cases, the level of the flat variance swap volatility has been calibrated such that the implied volatility of the three-year one-month ATM cliquet is 20%.

The values for  $\rho_{SX}$  and  $\rho_{SY}$  are  $\rho_{SX} = -70\%$  and  $\rho_{SY} = -35.7\%$  (with  $\chi = -50\%$ ). The corresponding term skew is that of figure 24.

In addition to the Black-Scholes price we discuss three other prices which are obtained by switching on either

- the one-month forward skew, i.e.

$$(\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%})_{1\text{ month}} \neq 0 \quad \text{and} \quad \omega = 0,$$

---

<sup>11</sup>A general cliquet option has the payoff

$$\min \left\{ \max \left\{ \sum_{i=1}^N \overline{r(t_i)}; F_g \right\}; C_g \right\},$$

where the the global floor  $F_g$  and global cap  $C_g$  are minimum and maximum returns the truncated (usual monthly) returns  $\overline{r(t_i)}$  are given by

$$\overline{r(t_i)} = \max \left\{ \min \left\{ \frac{S(t_i) - S(t_{i-1})}{S(t_{i-1})}; C_i \right\}; F_i \right\}$$

where the  $C_i$  and  $D_i$  are global caps and floors.  $\Delta = t_i - t_{i-1}$  is called the period of the cliquet.



- the volatility of volatility, i.e.

$$(\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%})_{1\text{ month}} = 0 \quad \text{and} \quad \omega \neq 0$$

- or both, i.e.

$$(\hat{\sigma}_{95\%} - \hat{\sigma}_{105\%})_{1\text{ month}} \neq 0 \quad \text{and} \quad \omega \neq 0.$$

In the next sections we give the definition of each product and comment on pricing results. The four corresponding prices are listed in table 3.

Model	Reverse Cliquet	Napoleon	Accumulator
Black-Scholes	0.25%	2.10%	1.90%
With Forward Skew	0.56%	2.13%	4.32%
With Vol of Vol	2.92%	4.71%	1.90%
Full	3.81%	4.45%	5.06%

Table 3: Prices of the Bergomi Model with constant Skew

### Reverse Cliquet

We consider a globally floored locally capped cliquet which pays once at maturity

$$\max \left\{ 0; C + \sum_{i=1}^N \min(r_i; 0) \right\}. \quad (4.5.1)$$

This structure is thereby termed a *reverse* cliquet because only negative returns contribute to the final payoff. Here the maturity is three years, the returns  $r_i$  are observed on a monthly basis ( $N = 36$ ) and the value of the coupon is  $C = 50\%$ .

From the values in table 3 we obtain that corrections to the Black-Scholes price are visible for the forward skew, the volatility of volatility and the full contribution and that the contribution of the volatility of volatility is by far the largest. The fact that the volatility of volatility contribution makes the reverse cliquet is more expensive is expected since a reverse cliquet has a structure similar to a Napoleon option for which we have deduced from figure 12 that it is in essence a put on volatility.

### Napoleon

We have still a maturity of three years and the option pays at the end of each year a coupon

$$\max \left\{ 0; C + \min_{i=1}^{12} r_i \right\}, \quad (4.5.2)$$

where  $\{r_i\}_{i=1,\dots,12}$  are the 12 monthly returns observed each year. Here we use  $C = 8\%$ . Again, the volatility of volatility had the biggest impact on the price, whereas the forward skew seems to have no impact.

### Accumulator

The maturity is again three years with one final payout given as a function of the 36 monthly returns  $\{r_i\}_{i=1,\dots,36}$ :

$$\max \left\{ 0; \sum_{i=1}^{36} \max (\min (r_i; \text{cap}) ; \text{floor}) \right\} ,$$

where  $\text{floor} = -1\%$  and  $\text{cap} = 1\%$  – a standard product.

Here the largest contribution comes from the forward skew. Note further that switching on the volatility of volatility in the case when there is no forward skew has no impact on the price while it does when the forward skew is switched on.

### The Effect of Changing the Spot/Volatility Correlation

In standard stochastic volatility models, changing the spot/volatility correlation changes the forward skew and thus the price of cliquets. In the Bergomi model, the spot process was specified such that it is homogeneous over the time interval  $[T_i, T_i + \Delta]$ , i.e. such that the distribution of  $\frac{S(T_i+\Delta)}{S(T_i)}$  does not depend on  $S(T_i)$ . Therefore changing the spot/volatility correlation does not change the value of one-month cliquets (ore in general of cliquets with a period of  $\Delta$ ). The change in the spot/volatility correlation only alters the *term* skew; the *short* forward skew is decoupled from the spot/volatility correlation, in particular the skew of maturity  $\Delta$  is independent of the spot level  $S(T_i)$ .

The prices listed in table 3 were calculated using  $\rho_{SX} = -70\%$  and  $\rho_{SY} = -35.7\%$ . Figure 24 shows that with these values the three-year 95-105% skew is 1.25%.

Now we halve the spot/volatility correlation:  $\rho_{SX} = -35\%$  and  $\rho_{SY} = -18\%$  ( $\chi = -19.2\%$ ). The three-year 95-105% skew is now 0.75%, also almost halved. The implied volatility of the three-year cliquet of the three-year cliquet of one-month ATM calls remains 20% and the price of a 95-105% one-month call spread cliquet is unchanged at 191.6%.

The new prices are shown in table 4. The difference with the prices in the fourth line of table 3 measures the impact of the *term* skew. The rest - in particular prices of cliquets - stays constant. The fact that the prices decrease when the spot/volatility correlation is less negative is consistent with the shape of the Black-Scholes vega as a function of the spot value (compare again figure 12).

Model	Reverse Cliquet	Napoleon	Accumulator
Full - Correlations halved	3.10%	4.01%	5.04%

Table 4: Prices of the Bergomi Model with constant Skew and Correlations halved

### Proportional short Skew

Here we discuss how a proportional model for the short skew alters prices using the three examples studied above. The functions  $\sigma_0 \left( \hat{\sigma}_{VS}^{(i)} \right)$  and  $\beta \left( \hat{\sigma}_{VS}^{(i)} \right)$  are calibrated such that the 95-105% skew for maturity  $\Delta$  is proportional to the ATMF volatility for maturity  $\Delta$  (remember also section 4.4.2). In particular the proportionality coefficient is calibrated such that the three-year cliquet of one-month 95-105% call spreads has the same value as before. The flat variance swap volatility is still chosen such that the implicit volatility of the three-year cliquet of a one-month

ATM call is 20%.

The corresponding prices are listed in table 5. We see that the accumulator and the reverse cliquet are now sizeable cheaper and that the price of the Napoleon does not change a lot.

Model	Reverse Cliquet	Napoleon	Accumulator
Full - Proportional Skew	3.05%	4.30%	4.15%

Table 5: Prices of the Bergomi Model with proportional Skew

## 4.6 Summary/Open Questions

In this chapter we introduced the Bergomi forward variance model. On the one hand this model has some remarkable advantages. It gives us the possibility to independently control

- the term-structure of the volatility of volatility,
- the short forward skew and
- the spot/volatility correlation.

Further this model is very flexible:

- It is able to incorporate any dependence of the short forward skew on the level of the at-the-money volatility.
- The contributions of the volatility of volatility and the forward skew can be turned on and off.
- The term skew can be changed while cliquet prices are kept constant.

On the other hand, there are some drawbacks and unsolved questions that are left for future work:

- Setting the parameters for the implied variance swap volatilities is a trading decision since by now there is no liquid market for caps, floors and swaptions on forward variances. Therefore it is not clear to what extent option prices depend on this step of the calibration.
- There is no explicit formula available for European calls and puts. Maybe there is the possibility of deriving a formula in a way similar to that of [Delbaen & Shirakawa, 2002], which we presented in section 2.3. But this seems to be a quite challenging task and it is possible that there is no closed-form solution at all.
- Even though the choice of the time scale  $\Delta$  is natural for many payouts - for example for Napoleons and reverse cliquets - it is more arbitrary for other options, for example for options on variance.

## References

- [Backus et al., 1997] Backus, D., Foresi, S., & Wu, L. (1997). Accounting for biases in black-scholes. Working Paper, New York University.
- [Bergomi, 2004] Bergomi, L. (2004). Smile dynamics. *Risk Magazine*, 17(9), 117–123.
- [Bergomi, 2005] Bergomi, L. (2005). Smile dynamics ii. *Risk Magazine*, 18(10), 67–73.
- [Bergomi, 2006] Bergomi, L. (2006). A model for consistent pricing of the smile dynamics. Presentation at Fraunhofer ITWM Kaiserslautern, October 2006.
- [Björk, 2004] Björk, T. (2004). *Arbitrage Theory in Continuous Time (Second Edition)*. Oxford University Press.
- [Brigo & Mercurio, 2006] Brigo, D. & Mercurio, F. (2006). *Interest Rate Models - Theory and Practice (Second Edition)*. Springer.
- [Brockhaus & Long, 2000] Brockhaus, O. & Long, D. (2000). Volatility swaps made simple. *Risk Magazine*, 13(1), 92–96.
- [Bronstein et al., 2000] Bronstein, I., Semendjajew, K., Musiol, G., & Mühlig, H. (2000). *Taschenbuch der Mathematik (5. Auflage)*. Verlag Harri Deutsch.
- [Cont, 2001] Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223–236.
- [Cox, 1975] Cox, J. C. (1975). Notes on option pricing i: Constant elasticity of variance diffusions. Working Paper, Stanford University, Graduate School of Business.
- [Delbaen & Shirakawa, 2002] Delbaen, F. & Shirakawa, H. (2002). A note of option pricing for constant elasticity of variance model. *Asia-Pacific Financial Markets*, 9(2), 85–99.
- [Dufresne et al., 2005] Dufresne, D., Garrido, J., & Morales, M. (2005). Fourier inversion formulas in option pricing and insurance. Working Paper.
- [Dupire, 1993] Dupire, B. (1993). Pricing and hedging with smiles. Proceedings of AFFI Conference, La Baule, June 1993.
- [Dupire, 1994] Dupire, B. (1994). Pricing with a smile. *Risk Magazine*, 7(1), 18–20.
- [Dupire, 2003] Dupire, B. (2003). Stochastic volatility modelling. Satellite Meeting to AMAM Colloquium, Nice, February 2003.
- [Franke et al., 2003] Franke, J., Härdle, W., & Hafner, J. (2003). *Einführung in die Statistik der Finanzmärkte (2. Auflage)*. Springer.
- [Gatheral, 2006] Gatheral, J. (2006). *The Volatility Surface: A Practitioner's Guide*. Wiley Finance.
- [Going-Jaeschke, 1998] Going-Jaeschke, A. (1998). *Parameter Estimation and Bessel Processes in Financial Models and Numerical Analysis in Hamiltonian Dynamics*. PhD thesis, ETH Zürich.
- [Heath & Schweizer, 2000] Heath, D. & Schweizer, M. (2000). Martingales versus pdes in finance: An equivalence result with examples. *Journal of Applied Probability*, 37, 947–957.

- [Heston, 1993] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2), 327–343.
- [Hull, 2005] Hull, J. C. (2005). *Options, Futures and other Derivatives (Sixth Edition)*. Pearson Prentice Hall.
- [Jacod & Protter, 2004] Jacod, J. & Protter, P. (2004). *Probability Essentials (Second Edition)*. Springer.
- [Javaheri, 2005] Javaheri, A. (2005). *Inside Volatility Arbitrage: The Secrets of Skewness*. Wiley Finance.
- [Jeffery, 2004] Jeffery, C. (2004). Reverse cliquets: End of the road? *Risk Magazine Review Article*, 17(2), 20–22.
- [Karatzas & Shreve, 2005] Karatzas, I. & Shreve, S. E. (2005). *Brownian Motion and Stochastic Calculus (Second Edition)*. Springer.
- [Korn, 2005] Korn, R. (2005). Probability theory 1. Lecture at TU Kaiserslautern, Winter Term 2004/05.
- [Korn, 2006] Korn, R. (2006). Four generations of equity price models. Presentation at Fraunhofer ITWM Kaiserslautern, October 2006.
- [Korn & Korn, 2001] Korn, R. & Korn, E. (2001). *Optionsbewertung und Portfolio Optimierung (2.Auflage)*. Vieweg Gabler.
- [Kraft, 2005] Kraft, H. (2005). Financial mathematics i: Stochastic calculus, option pricing, portfolio optimization. Lecture Notes, Summer Term 2005, TU Kaiserslautern.
- [Kraft, 2006] Kraft, H. (2006). Financial mathematics ii: Interest rate models. Lecture at TU Kaiserslautern, Winter Term 2005/2006.
- [Kruse & Nögel, 2005] Kruse, S. & Nögel, U. (2005). On the pricing of forward starting options in heston’s model on stochastic volatility. *Finance and Stochastics*, 9, 233–250.
- [Lemm, 2000] Lemm, J. C. (2000). Econophysics. Lecture Notes to the Lecture Econophysics, Winter Term 1999/2000, University of Münster.
- [Lewis, 2000] Lewis, A. L. (2000). *Option Valuation under Stochastic Volatility: With Mathematica Code*. Finance Press.
- [Mikhailov & Nögel, 2003] Mikhailov, S. & Nögel, U. (2003). Heston’s stochastic volatility model implementation, calibration and some extensions. *Wilmott Magazine*.
- [Neu-Ner, 2005] Neu-Ner, T. (2005). An effective binominal tree algorithm for the cev model. Honours Project, University of the Witwatersrand.
- [Primm, 2007] Primm, M. (2007). Das stochastische volatilitätsmodell nach heston. Diploma Thesis, TU Kaiserslautern.
- [Revuz & Yor, 1991] Revuz, D. & Yor, M. (1991). *Continuous Martingales and Brownian Motion*. Springer.
- [Schönbucher, 1999] Schönbucher, P. J. (1999). A market model for stochastic implied volatility. *Philosophical Transactions of the Royal Society*, 357(1758), 2071–2092.

- [Schroder, 1989] Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *The Journal of Finance*, 44(1), 211–219.
- [Swishchuk, 2004] Swishchuk, A. (2004). Modeling of variance and volatility swaps for financial markets with stochastic volatilities. *Wilmott Magazine*.
- [Tavella & Klopfer, 2001] Tavella, D. & Klopfer, W. (2001). Implying local volatility. *Wilmott Magazine*.
- [Wilmott, 2006] Wilmott, P. (2006). *Paul Wilmott on Quantitative Finance (Second Edition)*. John Wiley & Sons.
- [Yor, 1992] Yor, M. (1992). On some exponential functionals of brownian motion. *Advances in Applied Probability*, 24, 509–531.
- [Zhang, 2004] Zhang, A. (2004). The cox-ingersoll-ross process: Analytical properties and application in finance. Master’s Thesis, TU Kaiserslautern.